Racial Disparities in Infant Drug Testing
A Research Note

Mallory Dreyer, Cydney McGuire, Jennifer Silva,
Kosali Simon, Coady Wing, Engy Ziedan

Identifying maternal drug use during pregnancy can improve infant and maternal health by initiating treatment. However, drug tests may also be used for non-therapeutic purposes, such as criminal justice and child protective services involvement. Racial disparities in drug testing could undermine patient trust in the health care system, negatively impacting health outcomes. We used electronic medical records from $N = 140,562$ infants born between 2019-2023 in Indiana hospitals to measure meconium and umbilical cord tests for in-utero exposure to opioids and other illegal drugs. Black infants were tested at higher rates (16%) compared to White infants (7%), but both groups had a 20% positive testing rate. This could indicate racial bias if Black infants have a higher underlying risk of exposure, leading to over-testing of low-risk Black infants. We used inverse propensity score weights to decompose the testing gap into a portion explained by measured risk factors and an unexplained portion. Measured risk differences account for approximately 38% of the testing gap. The unexplained portion of the gap suggests that physicians use a different threshold for testing White and Black infants, or that physician decisions are motivated by additional risk factors that are not included in our adjustment.

Introduction

Substance use disorders are an important determinant of maternal and infant health outcomes in the United States (Jarlenski et al. 2020; Clemans-Cope et al. 2019). Identifying maternal drug use during pregnancy may improve maternal and infant health outcomes through pharmacological treatments, counseling, and postpartum services. However, infant and maternal drug tests are not used exclusively for medical decision-making. Drug test results also are sometimes reported to child protective services and criminal justice agencies. Although identifying substance use and child neglect are relevant to the missions of child protective services and criminal justice agencies, research suggests some pregnant women fear that a positive drug test puts them at risk for child separation and criminal justice system involvement (Roberts and Pies 2011). These fears may lead people to delay or avoid prenatal care, attempt a home birth, or attempt to detox without the supervision of a doctor (Stone 2015; Slaughter-Acey et al. 2019).

There is limited evidence on how often infant and maternal drug tests lead to non-medical consequences and on how these patterns differ by racial group. McCabe (2022) conducts qualitative interviews with health care providers and argues that providers seem to order tests for non-therapeutic reasons related to criminal justice concerns. In a retrospective chart review at a health system in Massachusetts, Cohen et al. (2023) found that about 4.8% of Black infants and 1.8% of White infants were tested for in-utero drug exposure. Among the infants who were drug tested, there was no racial disparity in the fraction who were referred to social work assessment or child protective services, or who experienced non-parental custody placement at discharge. Beyond the domain of infant and maternal health, research on physician expertise, predictive algorithms, and multistage assessment systems also suggests that there are racial disparities in medical treatment patterns that cannot be justified on purely clinical grounds (Currie and MacLeod 2017; Obermeyer et al. 2019; Baron et al. 2024).

Maternal and infant drug-testing practices are not standardized, and procedures for reporting drug tests to child protective agencies vary across states and across hospitals within states (Pregnant Women 2018; Mascola, Borders, and Terplan 2017). The American College of Obstetricians and Gynecologists (ACOG) recommends infant drug tests at birth only when physicians suspect exposure to harmful substances and a test is necessary for treatment. ACOG guidelines encourage physicians to assess an infant's risk of in-utero drug exposure in terms of the infant's health state (e.g. low birthweight or symptoms of withdrawal), and whether the mother's clinical record includes a history of drug misuse, late or inadequate prenatal care, and/or current tobacco or alcohol use (Palmer, Wood, and Krasowski 2017). Most previous work on racial disparities in infant and maternal drug testing is based on

relatively small samples collected from a single medical center or hospital network (Simpson et al. 2022; Soos et al. 2024; Schoneich et al. 2023; Kunins et al. 2007). This work suggests Black infants are tested at higher rates than White infants. However, existing research does not measure the degree to which racial differences in infant drug testing rates reflect differences in the clinical risk factors that are supposed to guide physician testing decisions.

In this study, we examine electronic medical records (EMR) on a large sample of 118,193 White and 22,369 Black newborn infants who were born between 2019 and 2023 at facilities that belong to the Indiana Network for Patient Care, which includes most of the hospitals and clinics in Indiana. For each infant, we observe whether a meconium or umbilical cord drug test was conducted and whether the test indicated in-utero drug exposure.[1] We link the records of infants and mothers and use them to measure clinically relevant risk factors.

We find that 16% of Black infants and 7% of White infants were tested. Among the infants who were tested, 20% of Black infants tested positive and 20% of White infants tested positive. At first glance, equal positive testing rates may suggest racial disparities in testing reflect differences in underlying risk rather than a bias towards testing Black infants. However, the fact that Black infants are tested at higher rates but have the same positive testing rate may also imply that physicians over-test low risk Black infants, bringing down the average positive testing rates. Without knowledge of the distribution of risk factors in the Black and White tested populations, these average positive testing rates alone do not reveal whether testing decisions involve racial prejudice or animus (Becker 1957; Knowles, Persico, and Todd 2001; Ayres 2002). The EMR data show that White and Black infants do have a different distribution of clinical risk factors. For example, mothers of Black infants had higher prevalence of maternal substance use history and resided in counties with higher rates of neonatal abstinence syndrome and drug overdose mortality. We use inverse propensity score weights to decompose the racial testing gap into a component explained by differences in clinically relevant risk factors and an unexplained component, which may reflect racial bias. The measured risk factors in our data account for approximately 38% of the Black-White testing gap. This implies that about 61% of the gap in testing arises because of differential treatment of infants with the same risk, or because clinicians may observe additional risk factors that are not captured in our data.

Data and Sample

The Indiana Network for Patient Care (INPC) harmonizes and combines electronic medical records (EMR) from multiple health systems, hospitals, and clinics in Indiana (Regenstrief

1. Meconium and umbilical cord infant drug tests measure maternal drug use in the second and third trimester of pregnancy (NMS Labs 2021).

Institute 2023). The EMR data contain patient demographics and healthcare encounter records that occur at contributing health care providers, including diagnosis codes, laboratory tests, and laboratory test results. Patients can be followed longitudinally using an encrypted person identifier.

Our study sample consists of infants born at INPC facilities between 2019 and 2023 who can be linked to a birth mother. We exclude infant-mother pairs with a residential zipcode outside of Indiana, maternal delivery age less than 12 or older than 60, missing maternal gender or maternal gender coded as male, or a discrepancy between infant date of birth and recorded date of delivery. We limit the sample to infants with with race coded as White or Black due to low cell sizes among the other racial/ethnic groups in our data.

Our final analytic sample contains one record for each of the $N = 140,562$ newborn infants who met inclusion criteria. The main outcome is a binary variable indicating whether a laboratory test for in-utero drug exposure was conducted on the infant. We include laboratory tests for opioids, benzodiazepines, cocaine, methamphetamine, amphetamines, and other drugs of abuse. Specifically, an infant is classified as tested if their medical record contains a lab test with a LOINC code corresponding to a family of meconium or umbilical cord tests for drugs of abuse – the full list of codes is in Appendix Table A1. A secondary outcome, measured only for infants who were tested, is a binary indicator that the test was positive for in-utero drug exposure.

In addition to lab test data, we also constructed a rich set of baseline covariates for each infant-mother pair. For the mother, we measure delivery age; the number of prenatal care visits; racial group; maternal insurance status (Medicaid, private, or other); and health history including prior diagnosis of chronic pain disorder, gestational diabetes, hypertensive disorder, mental health disorders, or substance use disorder. For the infant, we measure diagnosis of premature birth, low birth weight, and the number of healthcare encounters within one month of birth, which we view as a summary measure of infant health status.

We also attached two geographic measures of the severity of the opioid epidemic in the infant's home county. First, we used Centers for Disease Control (CDC) National Vital Statistics Mortality to compute the crude drug overdose mortality rate from 2014-2018 per 10,000 for each county in Indiana. Second, we used Indiana inpatient data from the Agency for Healthcare Research and Quality (AHRQ) Healthcare Cost and Utilization Project (HCUP) from 2018 to compute the rate of neonatal abstinence syndrome diagnosis per 100 births for each county in Indiana. We used the infant's home address to attach this county level information to each infant in our sample.

Methods

We compared testing rates between White and Black infants by year and within sub-populations defined by maternal Substance Use Disorder (SUD) history and the county prevalence of neonatal abstinence syndrome. We also examined positive test rates among tested infants by racial group and sub-population, which is one way to assess concerns that different clinical standards are used to decide whether to order a test for White and Black infants.

We use an inverse propensity score weighting strategy to decompose racial differences in testing rates into a component that reflects group differences in risk factors, and a component that may reflect structural differences in the way that health care providers respond to the same observed risk factors differently in White and Black infants (Fortin, Lemieux, and Firpo 2011).

To fix ideas, suppose X_i represents the set of clinically relevant factors that are observed by the clinician and are also contained in the EMR. U_i represents additional factors that may be observed by the clinician but are not available in the data. T_i is a binary variable indicating whether the clinician orders the test. To distinguish between racial group differences in testing rates that are due to clinical risks rather than differential treatment, assume test decisions are determined by $T_i = V_r(X_i, U_i)$, where $V_r(\cdot)$ is the physician's testing decision function – which could differ by race, $r = [W, B]$.

Under the assumption that unobserved clinical factors are independent of the infant's recorded race conditional on covariates, aggregate differences in testing rates across White and Black infants may arise for two reasons: (i) systematic differences in clinically relevant traits represented by X_i and U_i, and (ii) structural differences in the way that clinicians interpret clinically relevant traits for White and Black infants. In this notation, the overall difference in testing rates between White and Black infants is $\Delta_{B,W} = E[V_B(x,u)|B = 1] - E[V_W(x,u)|W = 1]$. This gap in testing rates can be decomposed into two pieces:

$$\Delta_{B,W} = \underbrace{E[V_B(x,u) - V_W(x,u) \mid W_i = 1]}_{\delta_s} + \underbrace{\big(E[V_B(x,u) \mid B_i = 1] - E[V_B(x,u) \mid W_i = 1]\big)}_{\delta_x}$$

In the formula, δ_s represents the difference in the testing rates that would prevail if Black infants had the same observable risk factors as White infants, but these risks were evaluated using race-specific decision functions. Conversely, δ_x represents the difference in testing rates that would be observed if White and Black infants possessed their actual observed risk factors but were both assessed using the Black decision function. Conceptually, δ_x is the

the part of the testing rate gap that arises because White and Black infants have different clinical risk factors. δ_s is the part of the observed racial testing gap that is not explained by observable differences in risk factors. It arises because clinicians interpret and respond to the same clinical information in different ways for White and Black infants, or because clinicians observe additional risk factors that are not recorded in the data. The share of the overall gap that is explained by risk factors is $\frac{\delta_x}{\Delta_{B,W}}$. The unexplained share is $\frac{\delta_s}{\Delta_{B,W}}$. We use an inverse propensity score weighting method to perform these decompositions and to estimate the structural and clinical component of racial testing gaps.[2]

Results

Infant Drug Testing by Race Across all years, 16.5% of Black infants are tested compared with 7.3% of White infants. The left panel of figure 1 shows that the testing gap was quite stable from 2019 to 2023. Despite the large gap in testing rates, the right panel of figure 1 shows that the positive testing rate was nearly identical among White and Black infants who were tested. The average positive testing rate for White infants was 19.7% compared to 20.2% among Black infants. The positive testing rate fell for both groups over time, particularly in 2022 and 2023.

Clinical Risk Differences Table 1 reports demographic and clinical characteristics for the mother and infant by race. Black infants were more likely to be born premature and to have low birth weight than White infants. Mothers of Black infants were about 1 year younger and received more prenatal care than mothers of White infants. Nearly 59% of the mothers of Black infants were covered by Medicaid compared with 26% of White mothers. About 12% of mothers of Black infants had a history of substance use disorder (SUD) and 6% had a history of a chronic pain disorder (CPD); among mothers of White infants, 10% had a previous SUD and 4% had a previous CPD diagnosis. In contrast, the mothers of Black infants had lower rates of mental health disorder than the mothers of White infants. The Black infants in our sample reside in counties with higher drug overdose mortality than White infants (2.7 vs 2.03 overdose deaths per 10,000). However, the county rates of neonatal abstinence syndrome (NAS) were about the same for both White and Black infants.

Testing Rates in Risk Sub-populations Figure 2 presents infant drug testing rates by race and maternal SUD history. For both White and Black infants, drug testing rates are about twice as high among those with maternal SUD history. In this high risk group, the

2. In the appendix, we also report estimates from Oaxaca-Blinder decompositions, which produces a similar decomposition using parametric regressions.

racial gap in testing rates has declined over time and is now quite small. Testing rates are lower among infants with no maternal SUD history, but the racial gap is substantial and more stable over time. Figure 3 illustrates that positive testing rates are about the same for White and Black infants, regardless of maternal SUD history.

The relationship between infant drug testing rates and neonatal abstinence syndrome (NAS) rates at the county level is depicted in Figure 4. Testing rates tend to be higher for both groups in counties that have higher rates of NAS, but the association is stronger for Black infants. Appendix Figure A1 shows that the positivity rate is essentially flat with respect to the NAS rate across counties. Appendix figures A2 and A3 show similar relationships with county drug mortality rates.

Decomposing Racial Testing Gaps We used inverse propensity score weights to reweight the White and Black infant samples to match the covariate distribution of the other group. Appendix table A2 reports the logistic regression we used to estimate propensity scores, and appendix table A3 shows covariate balance in the raw and reweighted sample.

Panel A of table 2 shows that the observed testing rate was 7.3% for White infants and 16.5% for Black infants. The second row shows implied testing rates when the data are reweighted. When the sample of Black infants is reweighted to match the covariate distribution of the White infants, the testing rate in the matched Black sample falls to 12.9%. When the White infants are reweighted to match the Black sample covariate distribution, the testing rate among White infants rises to 10.2%.

Figure 5 shows how testing rates vary over time among White infants, Black infants, and Black infants reweighted to match the White covariate distribution. Setting clinically relevant individual and community risk factors among Black infants equal to their levels in the White infant sample substantially narrows the testing gap but does not eliminate it.

Panel B of table 2 decomposes the racial testing gap into the share explained by observed risk factors and the share that is unexplained. The observed testing gap is $\Delta_{B,W} = 16.5\% - 7.3\%$, which is 9.2 percentage points. Reweighting the Black infants so that they have the same clinical risk distribution as the White sample but testing each group according to its own race-specific decision functions gives a structural gap of $\delta_s = 12.9\% - 7.3\% = 5.6$ percentage points. Applying the Black decision function to both White and Black infants leads to an risk-factor explained gap of $\delta_x = 16.5 - 7.3 = 3.5$ percentage points. This implies that clinical covariates account for about $100 \times 3.5/9.2 = 38\%$ of the raw testing gap. The remaining 61% of the gap may reflect either risk factors that are observed by health care workers but are not observed in the EMR data, or structural differences in the way that health care workers

interpret the same characteristics differently for White and Black mothers and infants. [3]

Discussion

Identifying mothers with substance use disorders during pregnancy may create an opportunity to provide medical care and social support that can improve maternal and infant health. But drug testing in medical settings is complicated. Maternal drug use is socially stigmatized, which may make discussions of drug testing uncomfortable and could make it difficult for health care providers to make objective decisions and recommendations. In some cases, drug testing may put the patient's medical best interests at odds with their best interests from a more holistic point of view, including legal or family stability implications. It is possible that some patients may refuse testing if they fear the test will be used against them. If they perceive medical staff to be untrustworthy, patients may cut back on medical care across the board. Furthermore, meconium and umbilical cord drug tests could be used to subject families to drug testing without their consent, raising further ethical dilemmas. These challenges are amplified by concerns that medical personnel use non-clinical factors like the patient's race to decide which patients to test for drug use during pregnancy.

This study examined drug testing decisions in a large study population of 140,562 newborns in Indiana. There was a large racial testing gap of 9.2 percentage points between Black and White infants, making the testing rate more than twice as high among Black infants compared with White infants. However, the positive testing rate among infants who were tested was about 20% for both Black and White infants. Despite the gap in testing rates, health care providers appear to allocate tests so that the average risk of substance use among those tested is the same across racial groups. This does not rule out the possibility that providers engage in discriminatory testing at the margin. Depending on the distribution of risks in the patient population, it is possible that physicians test "deeper" into the low risk end of the distribution for one racial group even if the overall average risk among the tested is the same.

Guidelines related to infant drug testing suggest that physicians should make testing decisions based on risks they observe from the health of the infant and the mother, including the medical history of the mother. In this paper, we used detailed EMR records to measure the main clinical risk factors that are supposed to guide physician testing decisions. Then we used an inverse propensity score weighting estimator to decompose the racial testing gap

3. Quantifying the share of the testing gap attributable to observable risks and unmeasured factors may be sensitive to which group is used as the reference category. In this case, the results are not very sensitive to the reference group: Panel C of table 2 shows that observed risk factors explain 33% of the gap when we change the reference group.

into a share that is accounted for by clinical risk factors and an unexplained component. We found that observed risk factors account for about 38% of the Black-White testing gap. In principle, the remaining 61% of the testing gap may reflect either risk factors that are observed by health care workers but are not observed in the EMR data, or differences in the way that health care workers interpret the same characteristics differently for White and Black samples. The fact that positive testing rates are about 20% for both White and Black infants may suggest that medical staff observe some additional risk factors that are predictive of substance use. Importantly, observing differences in risk profiles between Black and White racial groups should not be interpreted as evidence that racial identity itself carries risk (Chokshi, Foote, and Morse 2022). Differences in risk exposure could be rooted in upstream disparities in housing, employment, victimization, and education that contributes to poorer health behaviors and outcomes among historically marginalized groups (Yearby, Clark, and Figueroa 2022).

Our analysis focused on comparisons of realized testing status (tested or not) and on the test outcomes among the the tested. We did not study how patients were actually presented with the option to be tested or not tested, whether tested mothers explicitly consented to testing and knew that their infant was being tested, or whether untested mothers actually refused testing. These are important avenues for future research. It is possible – for example – that there are racial disparities in test refusal rates and in the way that a test refusal is respected or interpreted by medical personnel. We are not able to distinguish mothers who were presented with an opportunity to refuse or consent to infant testing, or to examine the post-decision health care utilization or outcomes of others who refused or consented to testing. Research on testing and consent procedures could help clarify whether and when drug testing undermines trust, and when it facilitates interventions that support the infant and mother.

Our paper has several limitations. First we cannot be sure that the electronic medical records contain all of the clinically relevant information underlying the test. It is possible that the physicians making testing decisions have additional information that justifies their decision and would explain the unexplained 61% of the gap in testing rates. Second, although we are able to compare the positive testing rate for the overall sample of Black and White infants, we are not able to measure positive testing rates for *marginally tested patients* in the sense used in the literature on *outcomes tests* for non-statistical discrimination (Ayres 2002; Knowles, Persico, and Todd 2001; Arnold, Dobbie, and Yang 2018). The challenge is that in our data, there is no quasi-experimental instrumental variable that shifts a patient's probability of being tested independently of their risk of testing positive. Research that compares positive testing rates at the margin may shed light on the possible role of racial

animus or false stereotypes in driving testing behavior. We hope to pursue these questions in future work using quasi-random variation in hospital staffing schedules to identify the marginally tested patients to better estimate the role of racial discrimination in rates of testing.

Despite these limitations, this study advances knowledge in several ways. We provide new evidence on clinician decisions on which infants to test for in-utero drug exposure using a large sample that covers a wide range of hospitals across Indiana. We document a large gap in testing rates and nearly identical positive testing rates. We also quantify the share of racial testing disparities that can be explained by differences in clinical risk factors. Our finding that significant disparities remain even after controlling for observable clinical risk factors calls for future investigations to better understand the process through which infants are tested, what clinicians do with the information they receive from test, and how mothers respond to and are affected by the test and corresponding outcomes.

References

Arnold, David, Will Dobbie, and Crystal S Yang. 2018. "Racial bias in bail decisions." *The Quarterly Journal of Economics* 133 (4): 1885–1932.

Ayres, Ian. 2002. "Outcome tests of racial disparities in police practices." *Justice research and Policy* 4 (1-2): 131–142.

Baron, E Jason, Joseph J Doyle Jr, Natalia Emanuel, Peter Hull, and Joseph Ryan. 2024. "Discrimination in Multiphase Systems: Evidence from Child Protection." *The Quarterly Journal of Economics,* qjae007.

Becker, Gary S. 1957. *The economics of discrimination.* University of Chicago press.

Chokshi, Dave A, Mary MK Foote, and Michelle E Morse. 2022. "How to act upon racism—not race—as a risk factor." In *JAMA Health Forum,* vol. 3, e220548–e220548. 2. American Medical Association.

Clemans-Cope, Lisa, Victoria Lynch, Embry Howell, Ian Hill, Nikhil Holla, Justin Morgan, Paul Johnson, Caitlin Cross-Barnet, and J Alice Thompson. 2019. "Pregnant women with opioid use disorder and their infants in three state Medicaid programs in 2013–2016." *Drug and alcohol dependence* 195:156–163.

Cohen, Samuel, Timothy Nielsen, Joseph H Chou, Bettina Hoeppner, Kathleen J Koenigs, Sarah N Bernstein, Nicole A Smith, Nicola Perlman, Leela Sarathy, Timothy Wilens, et al. 2023. "Disparities in maternal-infant drug testing, social work assessment, and custody at 5 hospitals." *Academic pediatrics* 23 (6): 1268–1275.

Currie, Janet, and W Bentley MacLeod. 2017. "Diagnosing expertise: Human capital, decision making, and performance among physicians." *Journal of labor economics* 35 (1): 1–43.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. "Decomposition methods in economics." In *Handbook of labor economics,* 4:1–102. Elsevier.

Jarlenski, Marian, Elizabeth E Krans, Qingwen Chen, Scott D Rothenberger, Abigail Cartus, Kara Zivin, and Lisa M Bodnar. 2020. "Substance use disorders and risk of severe maternal morbidity in the United States." *Drug and alcohol dependence* 216:108236.

Knowles, John, Nicola Persico, and Petra Todd. 2001. "Racial bias in motor vehicle searches: Theory and evidence." *Journal of political economy* 109 (1): 203–229.

Kunins, Hillary Veda, Eran Bellin, Cynthia Chazotte, Evelyn Du, and Julia Hope Arnsten. 2007. "The effect of race on provider decisions to test for illicit drug use in the peripartum setting." *Journal of women's health* 16 (2): 245–255.

Mascola, Maria A, Ann E Borders, and Mishka Terplan. 2017. "Opioid use and opioid use disorder in pregnancy." *Obstetrics and gynecology* 130 (2): 488–489.

McCabe, Katharine. 2022. "Criminalization of care: drug testing pregnant patients." *Journal of Health and Social Behavior* 63 (2): 162–176.

NMS Labs. 2021. *Newborn Toxicology Testing Comparison of Umbilical Cord Meconium and Urine.* Technical report. https://www.nmslabs.com/sites/default/files/2021-02/Newborn%20Toxicology%20Testing%20Comparison%20of%20Umbilical%20Cord%20Meconium%20and%20Urine.pdf.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. "Dissecting racial bias in an algorithm used to manage the health of populations." *Science* 366 (6464): 447–453.

Palmer, Kendra L, Kelly E Wood, and Matthew D Krasowski. 2017. "Evaluating a switch from meconium to umbilical cord tissue for newborn drug testing: a retrospective study at an academic medical center." *Clinical biochemistry* 50 (6): 255–261.

Pregnant Women, National Advocates for. 2018. *Understanding CAPTA and state obligations.*

Regenstrief Institute. 2023. *Guide to Data Sources, Version 1.2.* Available at www.ridata.org. Indianapolis, IN: Regenstrief Institute.

Roberts, Sarah CM, and Cheri Pies. 2011. "Complex calculations: how drug use during pregnancy becomes a barrier to prenatal care." *Maternal and child health journal* 15:333–341.

Schoneich, Sebastian, Melissa Plegue, Victoria Waidley, Katharine McCabe, Justine Wu, P Paul Chandanabhumma, Carol Shetty, Christopher J Frank, and Lauren Oshman. 2023. "Incidence of newborn drug testing and variations by birthing parent race and ethnicity before and after recreational cannabis legalization." *JAMA network open* 6 (3): e232058–e232058.

Simpson, Elizabeth A, David A Skoglund, Sarah E Stone, and Ashley K Sherman. 2022. "A prediction model for positive infant meconium and urine drug tests." *American journal of perinatology* 39 (10): 1104–1111.

Slaughter-Acey, Jaime C, Devon Sneed, Lauren Parker, Verna M Keith, Nora L Lee, and Dawn P Misra. 2019. "Skin tone matters: Racial microaggressions and delayed prenatal care." *American journal of preventive medicine* 57 (3): 321–329.

Soos, Alexandra, Melissa Plegue, Adam Darwiche, Lauren Oshman, and Christopher J Frank. 2024. "Racial Disparities in Newborn Drug Testing After Implementation of Question-Based Screening for Prenatal Substance Use." *Obstetrics Gynecology* 144 (2). ISSN: 0029-7844.

Stone, Rebecca. 2015. "Pregnant women and substance use: fear, stigma, and barriers to care." *Health & justice* 3:1–15.

Yearby, Ruqaiijah, Brietta Clark, and José F Figueroa. 2022. "Structural Racism In Historical And Modern US Health Care Policy: Study examines structural racism in historical and modern US health care policy." *Health Affairs* 41 (2): 187–194.

Tables and Figures

Table 1 — Maternal and Infant Characteristics by Racial Group of Infant

	Black Infants	White Infants
Count of Infants		
N	22,369	118,193
Maternal Characteristics		
Average Delivery Age	27.39	28.27
Age 15-19	0.00 (20)	0.00 (48)
Age 20-24	0.08 (1733)	0.05 (5623)
Age 25-29	0.28 (6210)	0.22 (25821)
Age 30-34	0.30 (6632)	0.32 (38090)
Age 35-39	0.22 (4814)	0.28 (32763)
Age 40-44	0.10 (2344)	0.11 (13216)
Age 45+	0.03 (616)	0.02 (2632)
Any Prenatal Care	0.68 (15321)	0.62 (72783)
12+ Prenatal Visits	0.24 (5465)	0.19 (22988)
Mom Black	0.80 (17937)	0.01 (638)
Mom Other	0.13 (2945)	0.07 (7801)
Mom White	0.07 (1487)	0.93 (109754)
Medicaid	0.59 (13108)	0.26 (31285)
Other Insurance	0.23 (5173)	0.37 (43647)
Private Insurance	0.18 (4088)	0.37 (43261)
Maternal Chronic Pain Disorder	0.06 (1305)	0.04 (4968)
Maternal Gestational Diabetes	0.07 (1673)	0.08 (9304)
Maternal Gestational Hypertensive Disorder	0.19 (4349)	0.16 (19014)
Maternal Mental Health Disorder	0.18 (4027)	0.21 (24860)
Maternal Substance Use Disorder	0.12 (2626)	0.10 (11573)
Infant Characteristics		
Average Crude Mortality Rate (2014-2018) per 10k	2.71	2.03
Average NAS Incidence per 100 Births	0.01	0.01
Average Number of Encounters Within 1 Month of Birth	1.70	1.59
Infant Tested	0.16 (3690)	0.07 (8655)
Infant Positive Test	0.03 (745)	0.01 (1707)
Low Birth Weight	0.06 (1283)	0.03 (3992)
Premature	0.06 (1425)	0.05 (5591)

Table 2 — Observed and Reweighted Means, and Decomposition Results

Panel A: Observed and Reweighted Means

Group	White	Black
Observed	0.0732	0.1650
Reweighted	0.1037	0.1295

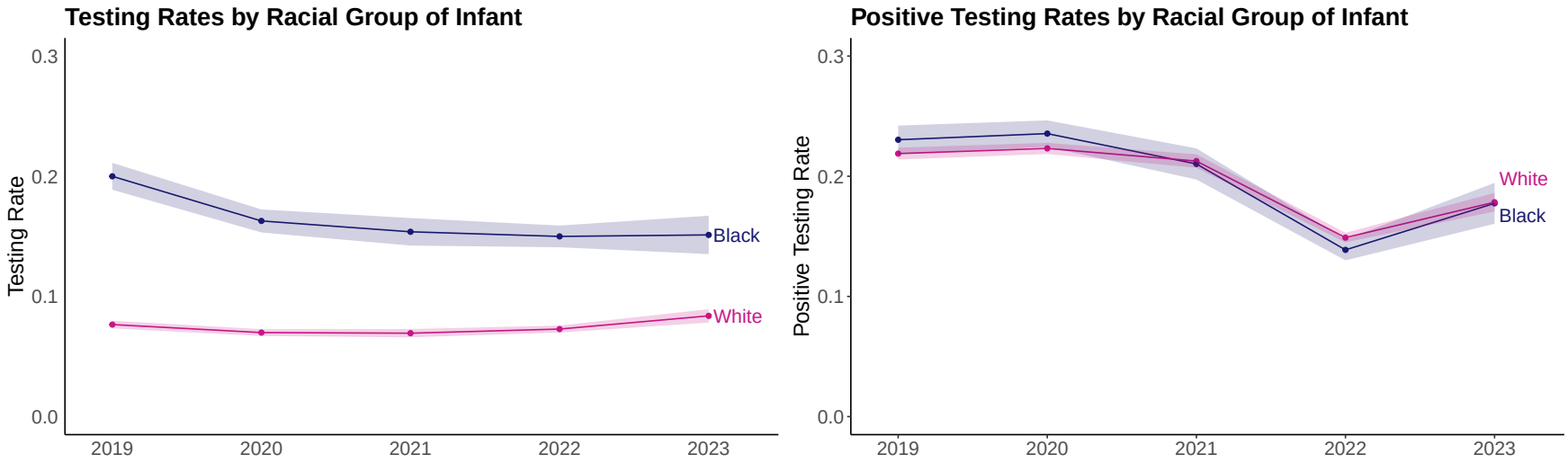
Panel B: Decomposition Results (Reweighting Black Infants)

Component	Gap	Standard Errors	Share of Raw Gap
Raw Gap	0.0917		1.0000
Reweighted Black - Observed White	0.0563	0.0037	0.6138
Observed Black - Reweighted Black	0.0354	0.0031	0.3862

Panel C: Decomposition Results (Reweighting White Infants)

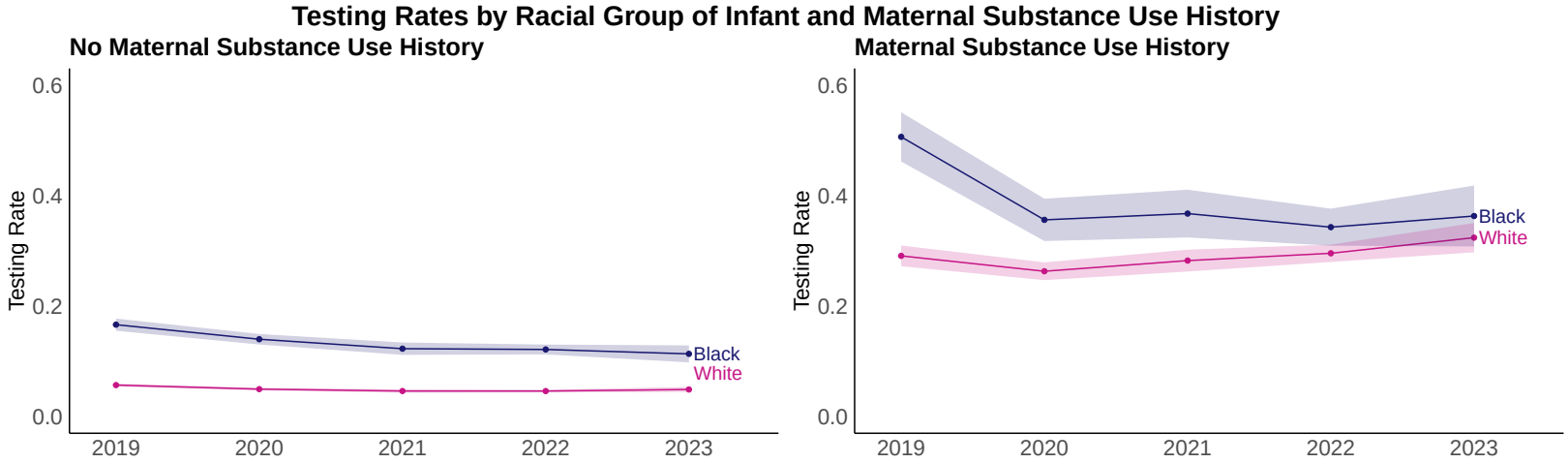
Component	Gap	Standard Errors	Share of Raw Gap
Raw Gap	0.0917		1.0000
Reweighted White - Observed White	0.0305	0.0012	0.3320
Observed Black - Reweighted White	0.0613	0.0029	0.6680

Figure 1 — Testing Rate and Positive Testing Rate by Racial Group of Infant by Year



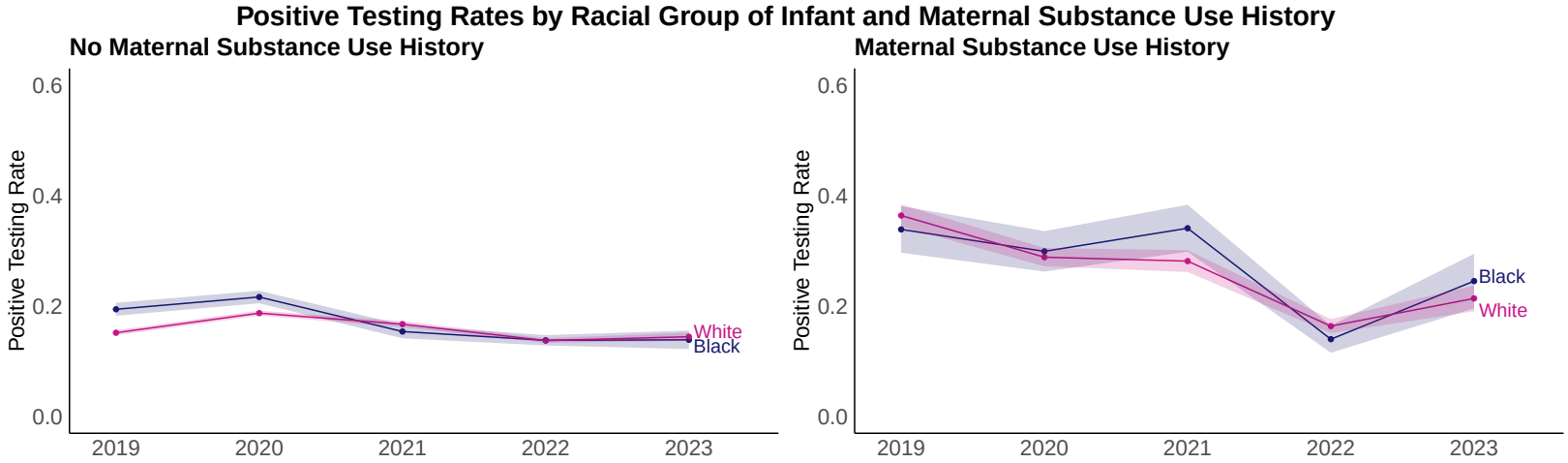
Note: The left panel shows the testing rate by year by racial group with the 95% confidence interval. The right panel shows the positive testing rate by year by racial group with the 95% confidence interval. The testing rate is calculated as the total number of infants tested divided by the total number of births in that year for each racial group. The positive testing rate is the number of infants with a positive test divided by the total number of infants tested in that year for each racial group.

Figure 2 — Testing Rates by Racial Group of Infant and Maternal Substance Use History



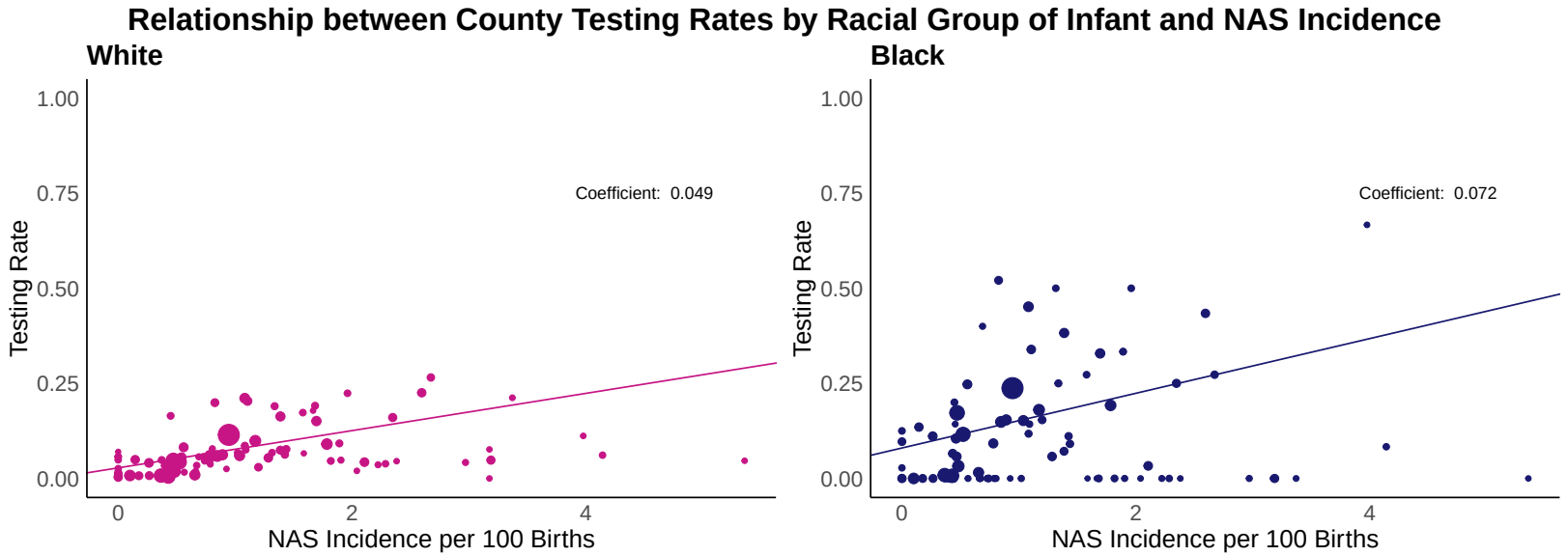
Note: The left panel shows the testing rate by year by racial group of infant with mothers without a history of substance use disorder (those without any previous diagnosis code for SUD in the EMR data). The right panel shows the testing rate by year and racial group of infants with mothers with a history of SUD. The testing rate is calculated as the number of infants tested in the given year, racial group, and maternal substance use history divided by the total number of infants born that year in the given racial group and maternal substance use history category.

Figure 3 — Positive Testing Rates by Racial Group of Infant and Maternal Substance Use History



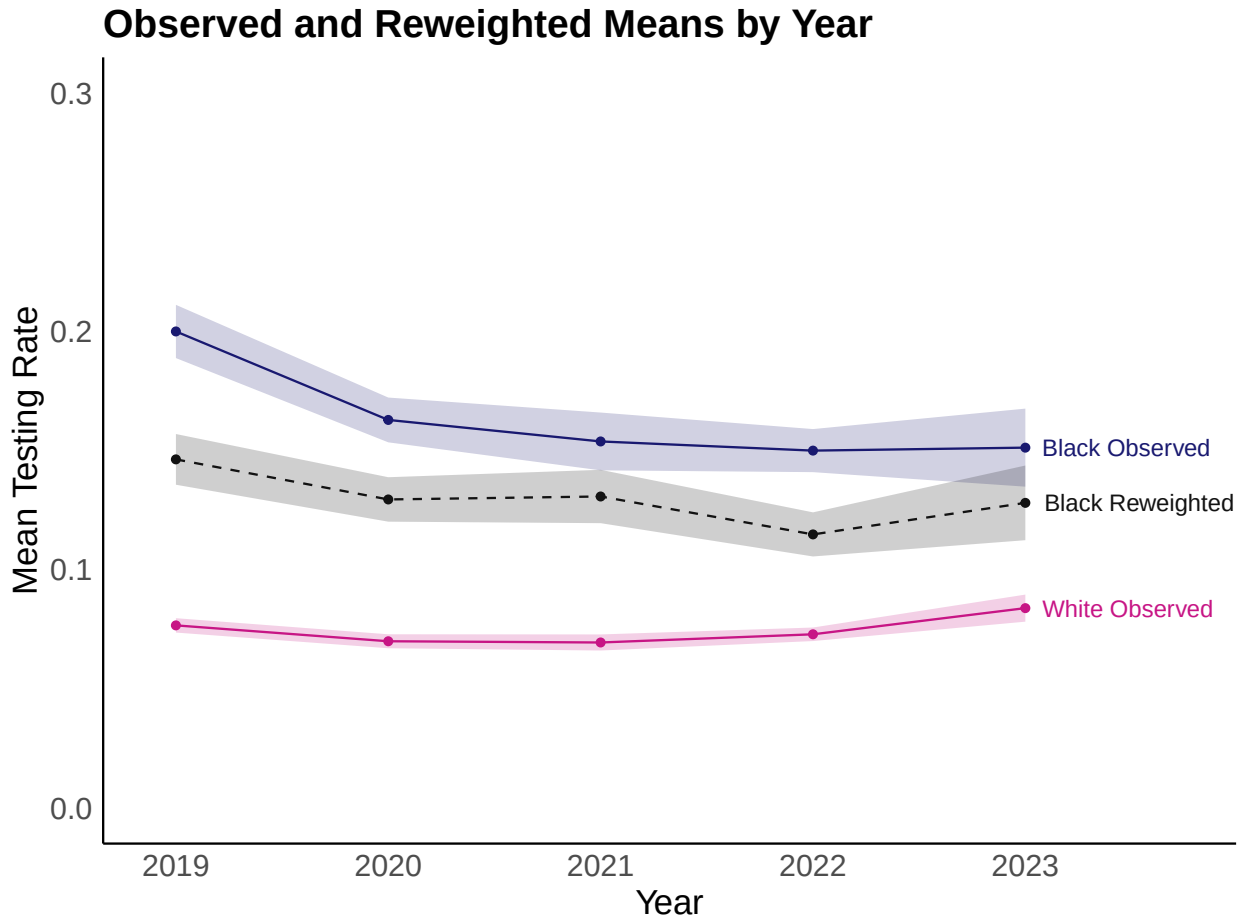
Note: The left panel shows the positive testing rate by year by racial group of infant for those with mothers without a history of substance use disorder (SUD). The right panel shows the positive testing rate by year by racial group of infant for those with mothers with a history of SUD. The positive testing rate is calculated as the number of infants with a positive test in the given year, racial group, and maternal substance use history category divided by the total number of infants tested for that group.

Figure 4 — Relationship between County NAS Incidence and County Average Testing Rate by Racial Group of Infant



Note: The left panel sows the relationship between the testing rate among White infants in the infant's county of residence and the county-level neonatal abstinence syndrome incidence per 100 births (all racial groups). Each point in the figure shows the average testing rate among White infants and the NAS incidence per 100 births in that county, and the size of the dot is weighted by the number of births in that county. The fitted line is a weighted linear regression with weighting based on the number of infants born in the county. The coefficient from the regression is 0.049 which means that an increase in the county-level NAS incidence of 1 is correlated with an increase in the county-level testing rate of 4.9 percentage points among White infants. Similarly, the right panel shows the relationship between the county-level NAS incidence per 100 births with the county-level testing rate of Black infants. The coefficient from the weighted linear regression of 0.072 which means that a 1 unit increase in the NAS incidence per 100 births is associated with a 7.2 percentage point increase in the testing rate of Black infants.

Figure 5 — Average Testing Rate by Racial Group of Infant and Reweighted Black Average for Decomposition



Note: This figure shows the average observed and reweighted testing rates by year to go along with the decomposition. The Black observed and White observed lines with the 95% confidence intervals are the same as Figure 1. The Black reweighted line shows what the average testing rate for Black infants would be if they looked like White infants. The line is lower than the Black observed line but does not overlap with the White observed line. Across all years, the difference between the observed Black and reweighted Black testing rates accounts for 38.6% of the difference in the observed Black and White testing rate. The difference between the reweighted Black and observed White average testing rates accounts for 61.4% of the overall gap.

Appendix

Lab Tests and LOINC Codes

The INPC EMR data includes lab tests and test results. For our sample of infants, we check to see if the infant had any meconium or umbilical cord drug tests performed. For those with tests performed, we check if any of the tests are positive or indicate the presence of the substance. The full list of LOINC codes for umbilical and meconium tests are in appendix table A1.

Propensity Score Model

We estimate propensity scores by fitting a logistic regression of the race of the infant on a vector of demographic, clinical, and community attributes that may be associated with the risk of in-utero drug exposure. Let W_i be a binary variable set to 1 if the infant is White and set to 0 if the infant is Black. Let X_i be the infant's vector of clinically relevant covariates. The basic model is:

$$Pr(W_i = 1|X_i) = logit^{-1}(X_i\beta)$$

The goal of the model is to produce inverse propensity score weights that balance the covariate distribution in the sample. We started with a parsimonious model in which all covariates were entered separately and maternal age was entered as dummy variables for 5-year bins (20-24, 25-29, etc.).

The final model is in Appendix Table A2. And the third column in Appendix Table A3 shows the reweighted means for the Black infants using the inverse propensity score weighting. The propensity score model is successful at forming a matched sample that is well balanced.

Estimates based on inverse propensity score weights may be unreliable when there is a failure of overlap or a near failure of overlap in the covariate distributions in the White and Black samples. In these cases, the inverse propensity score weights may be very large or very small. Figure A6 shows the distribution of weights in our sample and shows that no observation receives a weight that represents more than 0.28% of the sum of the weights in the sample. Figure A8 shows the binned distribution of estimated propensity scores in the White and Black samples. Both of these results indicate that our sample appears to have sufficient overlap and so we proceed without trimming.

Covid-19 Pandemic

Appendix Table A5 shows how the testing rate and the positive testing rate differ by the racial group of the infant before March 2020 and after March 2020. The testing rate for White infants remains relatively stable, while the testing rate for Black infants decreases by nearly 4 percentage points. The positive testing rate for both Black and White infants is lower in the period from March 2020 and after.

Using INPC COVID-19 lab test data, we construct a measure of the positive test rate from 2020-2023 in Indiana counties. In appendix figures A4 and A5, we show the correlation between the COVID-19 positive test rate and the test rate and positive test rate by racial group of infant by county for births occurring from March 2020 and after. For both Black and White infants, there is a negative correlation between the COVID-19 positive test rate and the in-utero drug exposure test rate, though the correlation is more negative for Black infants. Similarly, there is a negative correlation between the COVID-19 positive test rate and the in-utero drug exposure positive test rate for Black and White infants, with the Black infants having a more negative coefficient.

Table A1 — LOINC Codes for Umbilical and Meconium Drug Testing

Type	LOINC Codes
Umbilical	62364-5, 82373-2, 40626-4, 97242-2, 61042-8, 32080-4, 32081-2, 32088-7, 32093-7, 41859-0, 32099-4, 32100-0, 100357-3, 32101-8, 91053-9, 43811-9, 59355-8, 97306-5, 97306-5, 97292-7, 29530-3, 40609-0, 43230-2, 48946-8, 40625-6, 40481-4, 40381-6, 100358-1, 61038-6, 61037-8, 32057-2, 61039-4, 61031-1, 61074-1, 61044-4, 59712-0, 97278-6, 61051-9, 61055-0, 32108-3, 61061-8, 97310-7, 32107-5, 11526-1, 82375-7, 97286-9, 97290-1, 97296-8, 93121-2
Meconium	19125-4, 73728-8, 26911-8, 15403-9, 26914-2, 15402-1, 26979-5, 26815-1, 57945-8, 40418-6, 26938-1, 11023-9, 29160-9, 40414-5, 27003-3, 15407-0, 27302-9, 27029-8, 26937-3, 26913-4, 27324-3, 43933-1, 26934-0, 26727-8, 59630-4, 29161-7, 40412-9, 26716-1, 27004-1, 26935-7, 27035-5, 26863-1, 57981-3, 26862-3, 57982-1, 26936-5, 26857-3, 15405-4, 27289-8, 29286-2, 11026-2, 27262-5, 27244-3, 27080-1, 26866-4, 40385-7, 27308-6, 8232-1, 26712-0, 27235-1, 26713-8, 27002-5, 26893-8, 8167-9, 59588-4, 26959-7, 8143-0, 26895-3, 8144-8, 27061-1, 27962-0, 27024-9, 60519-6, 15364-3, 26865-6, 29278-9, 11022-1, 15370-0, 26860-7, 29287-0, 26933-2, 26869-8, 27274-0, 31081-3, 27026-4, 8187-7, 32797-3, 58738-6, 49046-6, 26912-6, 11025-4, 29159-1, 43595-8, 8213-1, 29158-3, 8214-9, 8231-3, 26859-9, 26714-6, 8166-1, 40901-1, 59610-6, 73730-4, 30108-5, 76349-0, 69011-5, 26956-3, 40527-4, 34578-5, 38441-2, 8186-9, 64126-6, 75647-8, 59749-2, 93719-3, 26910-0, 8146-3, 69021-4, 31080-5, 38439-6, 59621-3, 73996-1, 8189-3, 59686-6, 69022-2, 69024-8, 69025-5, 59650-2, 26891-2, 20481-8, 26744-3, 8216-4, 69026-3, 26829-2, 8234-7, 11027-0, 27067-8, 8168-7, 8169-5, 93720-1, 29345-6, 43934-9, 8145-5, 18278-2, 33282-5, 50406-8, 94399-3, 59631-2, 69008-1, 68542-0, 69027-1, 59729-4, 68543-8, 69028-9, 20480-0, 8188-5, 59658-5, 68544-6, 68541-2, 91030-7, 91031-5, 69023-0, 93722-7, 75379-8, 8215-6, 50620-4, 92816-8, 8233-9, 26711-2, 69012-3, 73563-9, 77205-3, 94398-5, 69009-9, 41037-3, 91032-3, 69010-7, 59694-0, 73567-0, 73562-1, 59603-1, 73564-7, 75367-3, 59720-3, 93721-9, 31136-5, 57924-3, 91029-9, 69007-3, 15406-2, 15365-0, 15371-8

Table A2 — Propensity Score Model

	(1)
Intercept	-4.8949***
	(0.5059)
Age 20-24	0.0513
	(0.5059)
Age 25-29	-0.0225
	(0.5035)
Age 30-34	-0.2443
	(0.5035)
Age 35-39	-0.3588
	(0.5039)
Age 40-44	-0.3345
	(0.5053)
Age 45-49	-0.1561
	(0.5124)
Above Average NAS in County	-0.5477***
	(0.0448)
Any Maternal SUD	0.2283***
	(0.0488)
Any Prenatal Care	0.0722*
	(0.0349)
Above Average Overdose Mortality in County	0.8008***
	(0.0341)
Any Maternal Mental Health Diagnosis	-0.1399***
	(0.0407)
Any Maternal Hypertension	0.2157***
	(0.0419)
Medicaid	0.9602***
	(0.0420)
Other Insurance	0.1328**
	(0.0460)
Low Birthweight	0.4080***
	(0.0799)
Premature Birth	-0.0884
	(0.0749)
Four or More Encounters	0.0656
	(0.0467)
Mom Black	7.3683***
	(0.0487)
Mom Other	2.9920***
	(0.0352)
Num.Obs.	140562

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A3 — Maternal and Infant Characteristics by Racial Group of Infant with Reweighted Black Sample

	Black Infants	White Infants	Black Infants Reweighte
Total Number of Infants			
Count	22369	118193	
Maternal Characteristics			
Average Delivery Age	27.39	28.27	28.17
Age 15-19	0.00 (20)	0.00 (48)	0.00
Age 20-24	0.08 (1733)	0.05 (5623)	0.05
Age 25-29	0.28 (6210)	0.22 (25821)	0.23
Age 30-34	0.30 (6632)	0.32 (38090)	0.31
Age 35-39	0.22 (4814)	0.28 (32763)	0.27
Age 40-44	0.10 (2344)	0.11 (13216)	0.11
Age 45+	0.03 (616)	0.02 (2632)	0.02
Any Prenatal Care	0.68 (15321)	0.62 (72783)	0.60
12+ Prenatal Visits	0.24 (5465)	0.19 (22988)	0.21
Mom Black	0.80 (17937)	0.01 (638)	0.78
Mom Other	0.13 (2945)	0.07 (7801)	0.12
Mom White	0.07 (1487)	0.93 (109754)	0.11
Medicaid	0.59 (13108)	0.26 (31285)	0.26
Other Insurance	0.23 (5173)	0.37 (43647)	0.36
Private Insurance	0.18 (4088)	0.37 (43261)	0.38
Maternal Chronic Pain Disorder	0.06 (1305)	0.04 (4968)	0.05
Maternal Gestational Diabetes	0.07 (1673)	0.08 (9304)	0.08
Maternal Gestational Hypertensive Disorder	0.19 (4349)	0.16 (19014)	0.16
Maternal Mental Health Disorder	0.18 (4027)	0.21 (24860)	0.22
Maternal Substance Use Disorder	0.12 (2626)	0.10 (11573)	0.11
Infant Characteristics			
Crude Drug Mortality Rate	2.71	2.03	2.12
Average NAS Incidence	0.01	0.01	0.01
Encounters Within 1 Month of Birth	1.70	1.59	1.62
Infant Tested	0.16 (3690)	0.07 (8655)	0.13
Infant Positive Test	0.03 (745)	0.01 (1707)	0.02
Low Birth Weight	0.06 (1283)	0.03 (3992)	0.03
Premature	0.06 (1425)	0.05 (5591)	0.05

Table A4 — Oaxaca Decomposition Results with Calculation Details

Panel A: Decomposition Results (Black as Baseline)

Component	Calculation	Gap	Standard Errors	Share of Raw Gap
Raw Gap	$\bar{X}_{Black} - \bar{X}_{White}$	0.0917		1.0000
Endowments	$\Delta\bar{X} \cdot \beta_{Black}$	0.0375	0.0027	0.4088
Coefficients	$\Delta\beta \cdot \bar{X}_{White}$	0.0542	0.0032	0.5912

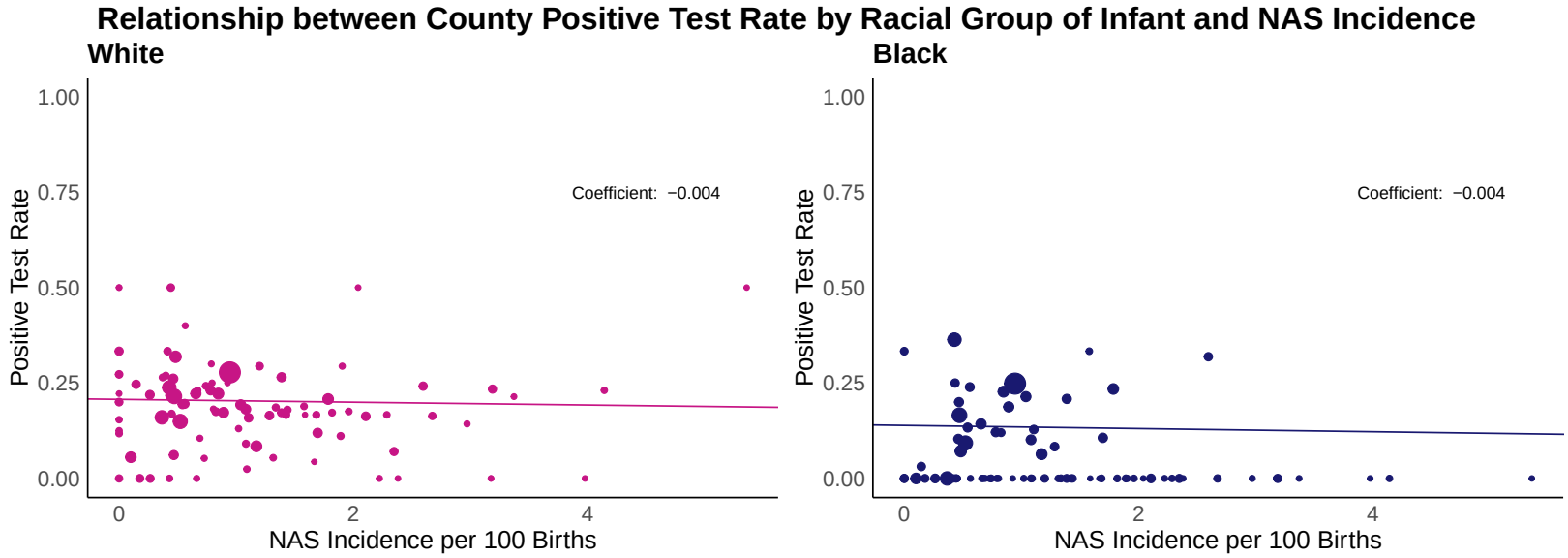
Panel B: Decomposition Results (White as Baseline)

Component	Calculation	Gap	Standard Errors	Share of Raw Gap
Raw Gap	$\bar{X}_{Black} - \bar{X}_{White}$	0.0917		1.0000
Endowments	$\Delta\bar{X} \cdot \beta_{White}$	0.0302	0.0011	0.3289
Coefficients	$\Delta\beta \cdot \bar{X}_{Black}$	0.0616	0.0027	0.6711

Table A5 — Infant Drug Testing Rates and Positive Rates by Racial Group During the COVID-19 Pandemic

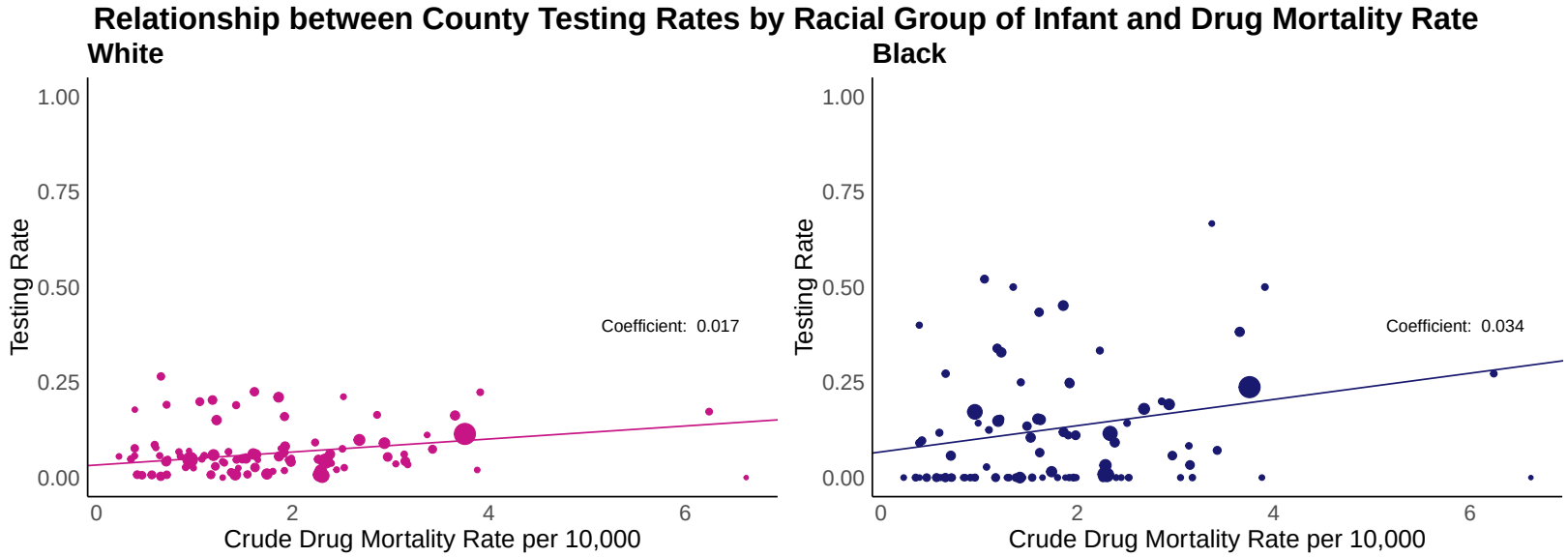
	Before March 2020		March 2020 and After	
Race	Testing Rate	Positive Rate	Testing Rate	Positive Rate
Black	0.1917	0.2310	0.1554	0.1891
White	0.0756	0.2183	0.0723	0.1890

Figure A1 — Relationship between County NAS Incidence and County Average Positive Testing Rate by Racial Group of Infant



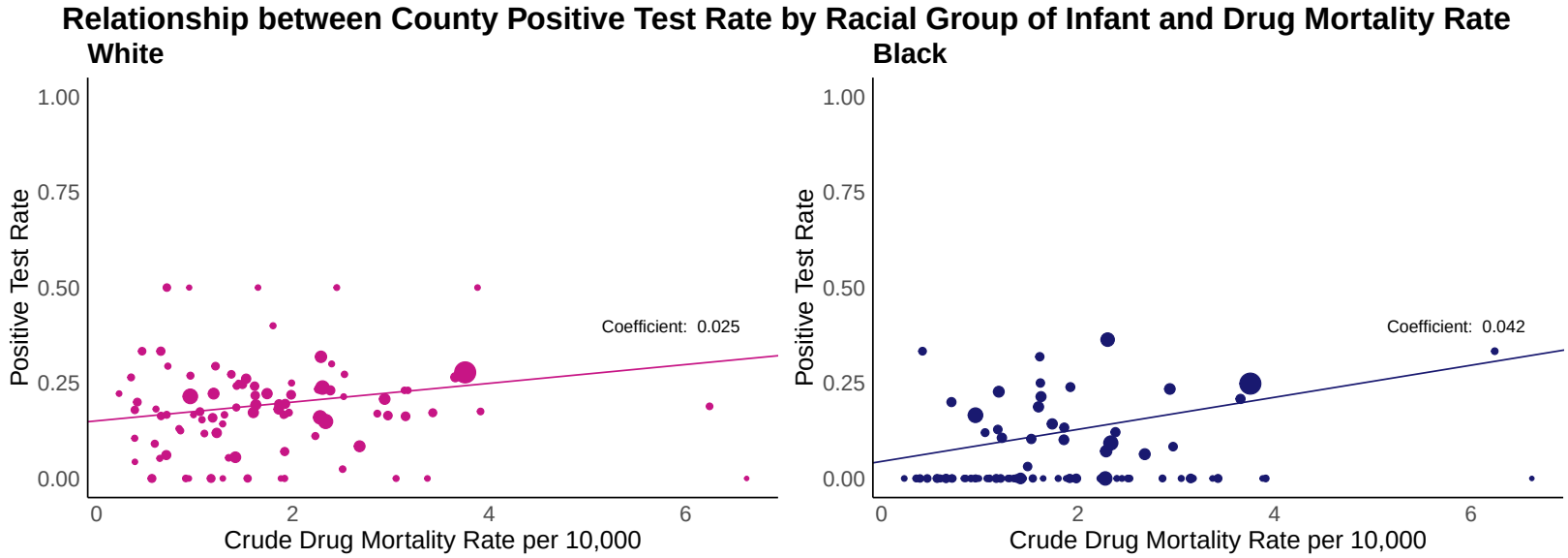
Note:

Figure A2 — Relationship between County Crude Drug Mortality and County Average Testing Rate by Racial Group of Infant



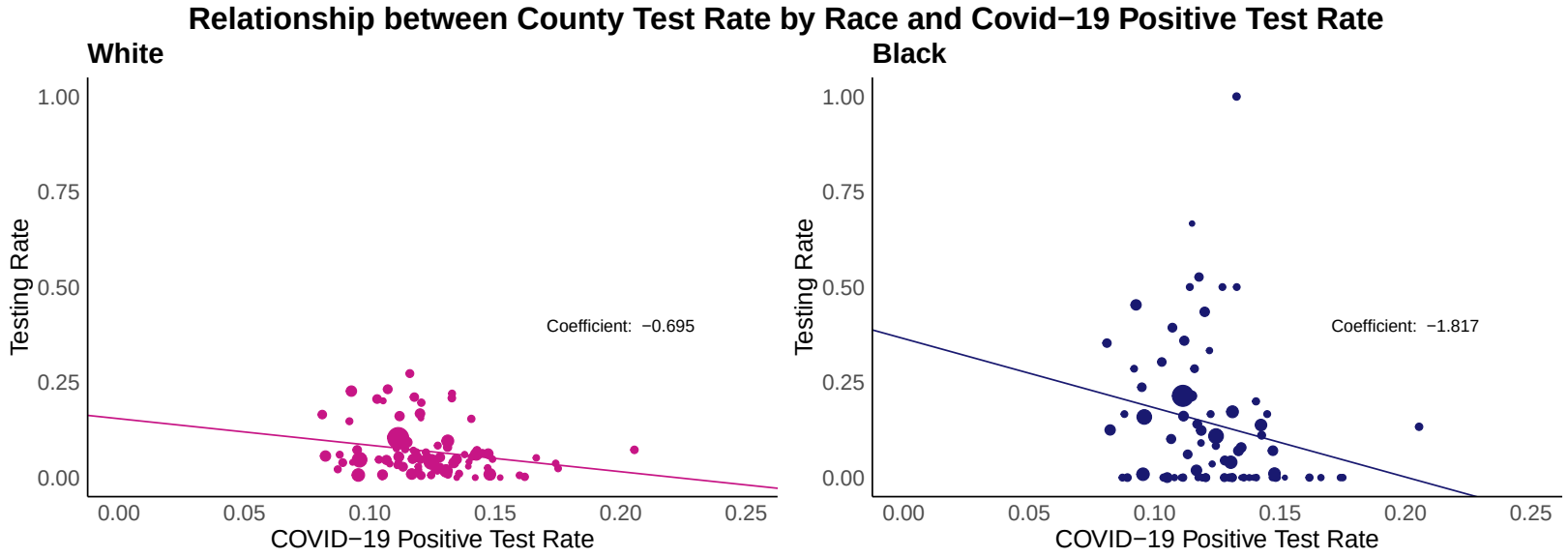
Note:

Figure A3 — Relationship between County Crude Drug Mortality and County Average Positive Testing Rate by Racial Group of Infant



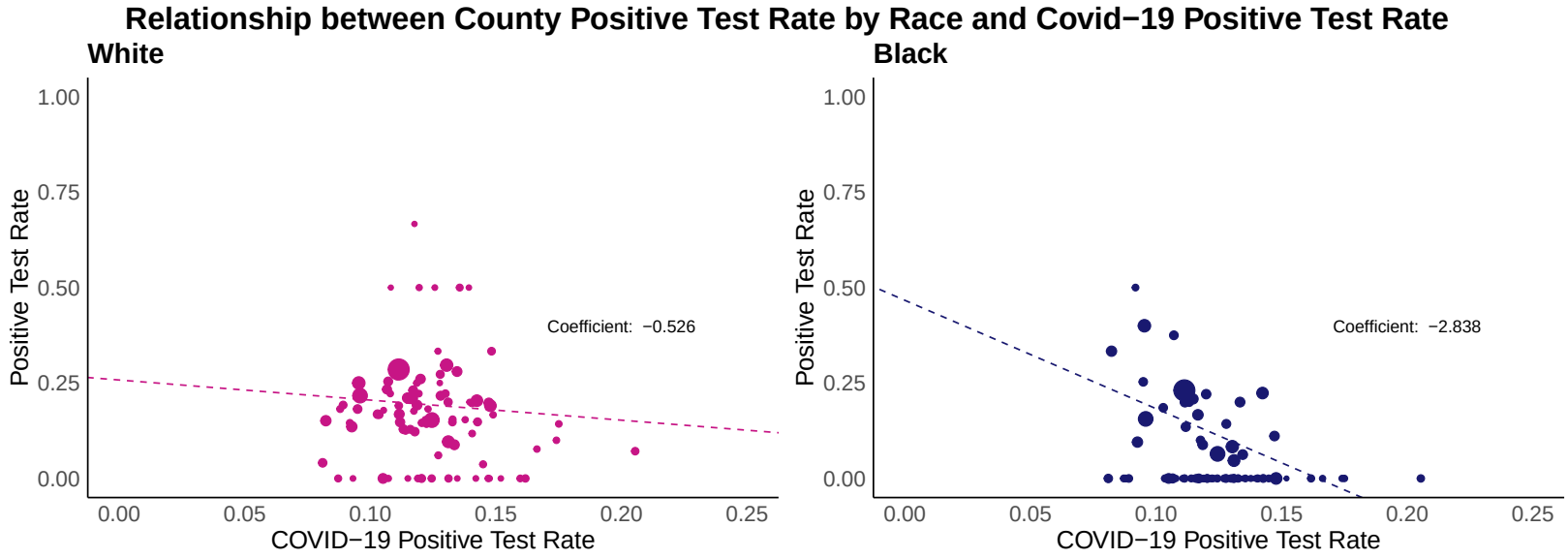
Note:

Figure A4 — Relationship between County COVID-19 Positive Test Rate and County Average Testing Rate by Racial Group of Infant



Note:

Figure A5 — Relationship between County COVID-19 Positive Test Rate and County Average Positive Testing Rate by Racial Group of Infant



Note:

Figure A6 — Distribution of the Fraction of Total Weight Amongst Black Infants Reweighted to be Like White

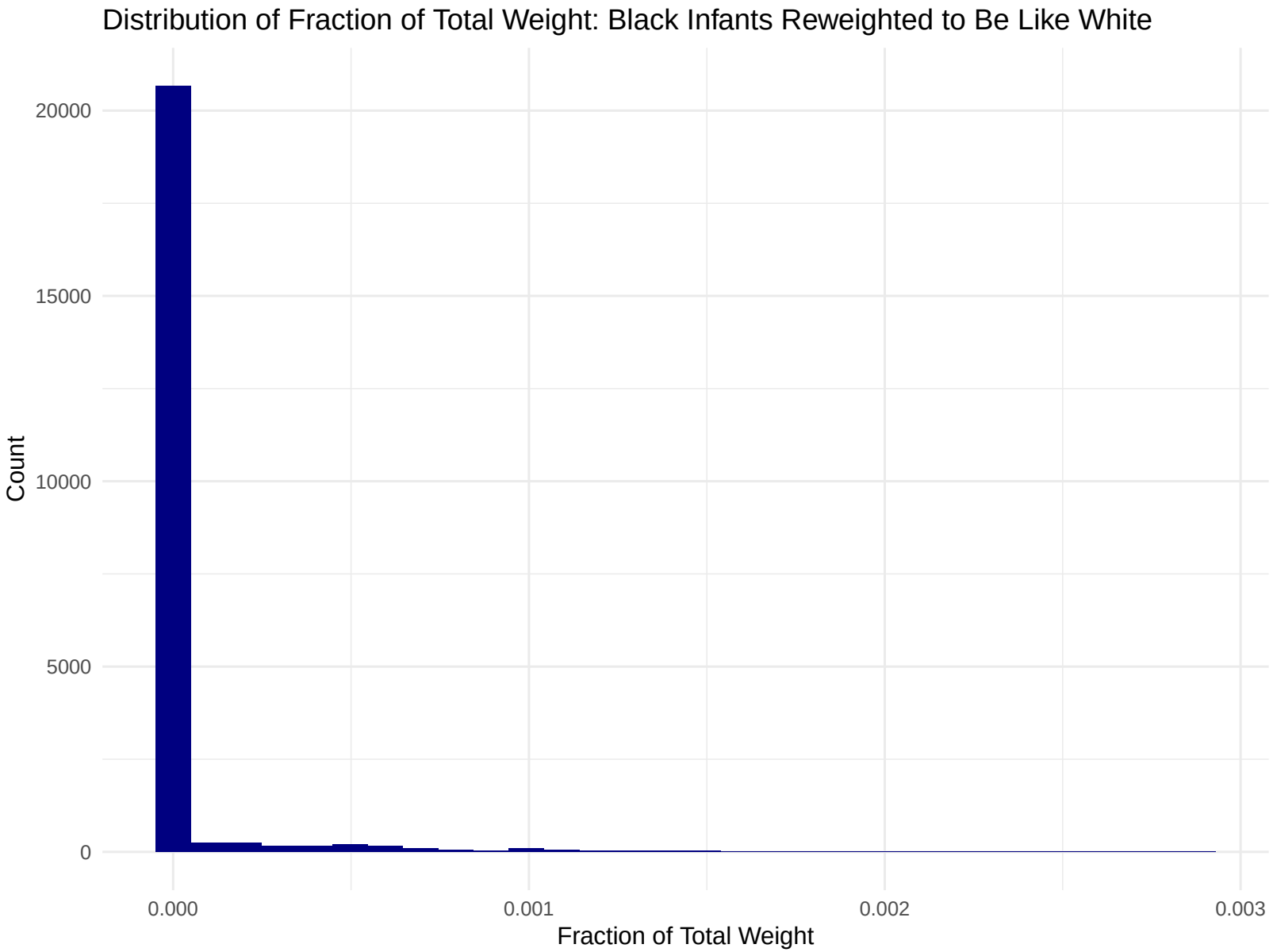


Figure A7 — Overlap in Propensity Scores by Racial Group of Infant

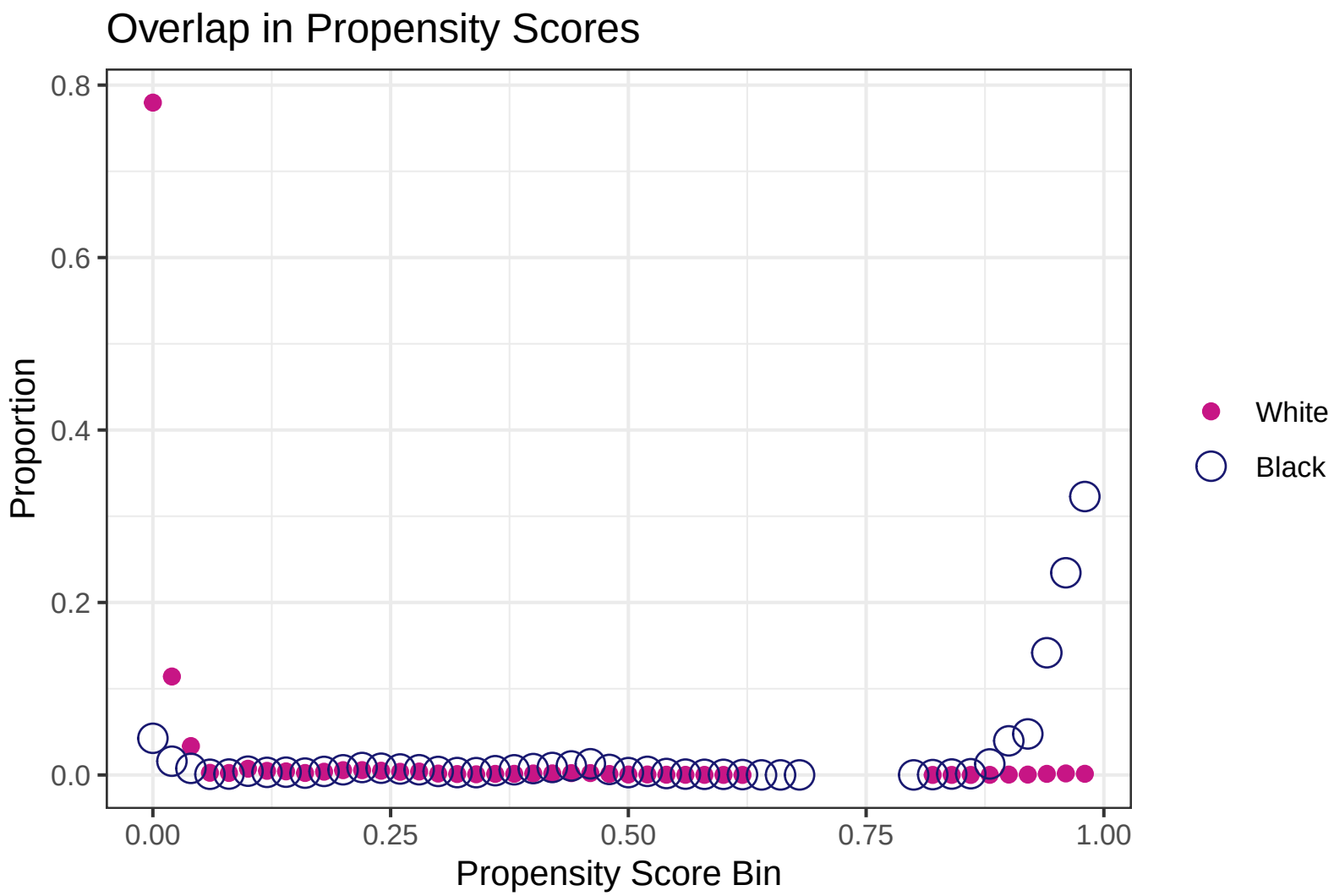


Figure A8 — Infant Drug Testing Rates by Racial Group of Infant and Maternal Drug Testing Status during Pregnancy

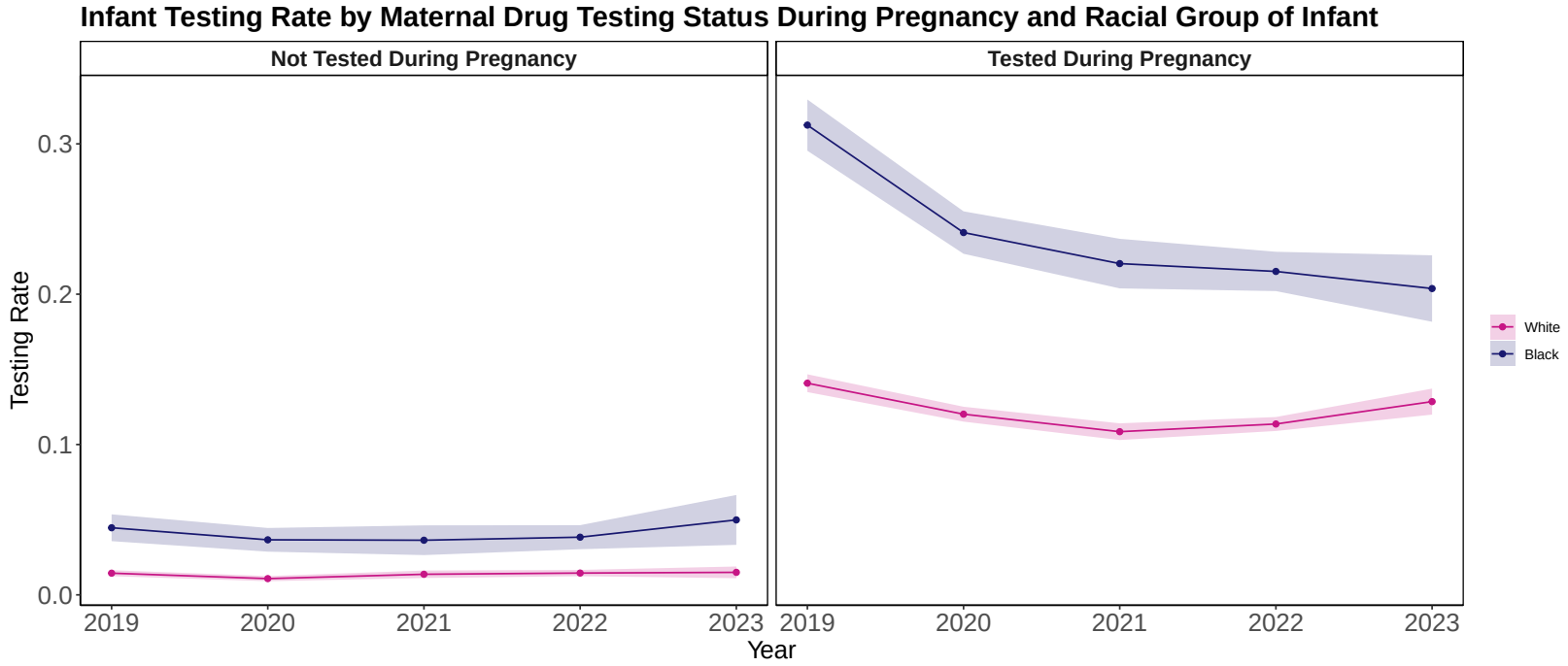


Table A6 — Comparison of Analytic Sample from INPC with All Births in Indiana from the National Vital Statistics System

Variable	INPC	NVSS
N	140,562	320,601
Age	28.67	28.19
Age 15-19	0.00	0.05
Age 20-24	0.05	0.23
Age 25-29	0.23	0.32
Age 30-34	0.32	0.27
Age 35-39	0.27	0.11
Age 40-44	0.11	0.02
Age 45+	0.02	0.00
Any Prenatal Care	0.63	0.98
12+ Prenatal Visits	0.20	0.51
Mom Black	0.13	0.13
Mom White	0.79	0.82
Mom Other	0.08	0.05
Medicaid	0.32	0.39
Other Insurance	0.35	0.09
Private Insurance	0.34	0.52
Gestational Diabetes	0.08	0.08
Gestational Hypertension	0.17	0.10
Premature Birth	0.05	0.11
LBW	0.04	0.08

Note: This table compares the full analytic sample from INPC with all births in Indiana from the National Vital Statistics System. The births in INPC are limited to those infants whose racial group is either Black or White. The NVSS births are all births, as there is no variable for the race of the infant–only the race of the mother is listed.

Table A7 — Positive Testing Rates Among Infants

Panel A: Reweighted Based on Full Sample Propensity Scores		
Group	**White**	**Black**
Observed	0.1972	0.2019
Reweighted	0.1977	0.1388
Panel B: Reweighted Based on Only Tested Sample Propensity Scores		
Group	**White**	**Black**
Observed	0.1972	0.2019
Reweighted	0.1625	0.1289