# The Impact of Intensive Nurse Home Visiting on Pregnancy and Postpartum Mental Health for Low-Income First-Time Mothers [*]

Abigail Dow[1‡]        Dea Oviedo [2]

Michelle Woodford Martin[3]        Mary Ann Bates[4]

Katherine Baicker[5]        Alecia McGregor[2]

Margaret McConnell[2]

July 21, 2025

**Abstract**

Depression during and after pregnancy is more common for low income women, with nearly a quarter experiencing depression across the perinatal period. Intensive supplemental programs, such as nurse home visiting, have been proposed as a way to improve mental health outcomes for low-income women and address racial disparities originating from social determinants of health. These racial disparities are evident in mental health outcomes; Black women are less likely to have their depression diagnosed or treated. We evaluate the impact of the Nurse Family Partnership (NFP) on mental health outcomes and racial disparities using a large-scale RCT in South Carolina. We find that NFP had no statistically significant impacts on mental health diagnoses and treatment overall. Results show that NFP did not reduce racial disparities in outcomes; NFP participation lowered depression treatment rates for Black mothers but increased rates for White mothers, though differences were not robust to multiple inference adjustment.

# 1  Introduction

Depression during pregnancy or the postpartum period affects many women - 17% of mothers experience antenatal depression and 13% of mothers experience postpartum depression (Dagher et al., 2021). The negative relationship between income and depression (Ettner, 1996; Gresenz et al., 2001; Zimmerman and Katon, 2005) documented in the general populations is also present amongst mothers. Postpartum depression affects one in four low-income women (Kozhimannil and Kim, 2014), and similar rates of Black women (Getahun et al., 2023). Maternal depression can have potentially severe consequences for the mother and infant. Left untreated, it can increase the risk of suicide, substance abuse, and missed health care appointments for the mother, and for the child, it increases the risk of adverse health outcomes (Casey et al., 2004; Perry, 2008; Propper et al., 2007), development issues (Deave et al., 2008), behavioural problems (Frank and Meara, 2009) and is linked to worse later life outcomes (Johnston et al., 2013). Moreover, the consequences of maternal depression are increasingly recognised (Kim, 2024) as potential drivers of racial and economic inequalities in maternal and infant health (Kennedy-Moulton et al., 2022; Case et al., 2002).

Policymakers have invested in interventions targeting low-income mothers to try to address health inequalities. Pregnancy and early childhood interventions, in particular, have been shown to have positive effects on health, education, and earnings throughout the life course (e.g. Currie and Almond (2011); Heckman et al. (2013); Conti et al. (2016); Campbell et al. (2014)). This paper explores the effects of one such intervention: intensive nurse home visiting.

We look at the impact of intensive nurse home visiting on perinatal mental health utilisation outcomes for a low-income, Medicaid eligible population. Intensive home visiting programs are designed to develop a long-term relationship between a home visitor and a pregnant person during the transition to parenthood. They have the potential to reduce anxiety and stress in the perinatal period by providing mothers with support, advice, and reassurance. These programs may also increase the identification and treatment of maternal mental health problems. Furthermore, home visiting programs often seek to support new mothers with challenges that are closely linked with maternal mental health, including substance abuse and conflict between partners.

In this paper, we evaluate the effects of an intensive nurse home visiting program on maternal mental health diagnoses and treatment, and substance use and domestic violence, leveraging a large-scale randomised controlled trial of the Nurse-Family Partnership (NFP) in South Carolina that enrolled between 2016 and 2020, and followed outcomes through 2022. We look at outcomes during pregnancy and up to sixty days postpartum using multiple linked datasets (including administrative records): survey data collected from mothers at baseline, NFP program data, vital records data, Medicaid claims and hospital discharge data, and South Carolina Department of Mental Health data. We find that NFP has no statistically significant impact on mental health diagnoses and treatment, nor on substance use or experiences of violence.

The overall finding of null effects of NFP in the full sample could be explained by multiple factors. Mothers suffering from depression and anxiety may struggle to engage with these intensive programs and have been found to exit home visiting programs if they lack access to outside mental health services (Molina et al., 2020). Additionally, a survey of NFP nurses reveals that nurses report lacking confidence in evaluating mothers' mental health and meeting their needs (Beeber et al., 2022). For people with mental health conditions, there may also be particular challenges in implementing intensive home visiting programs. The programs cannot be implemented without nurses having consistent and prolonged access to mothers in their homes (Duggan et al., 2018), which mothers could prevent.

Prior work has shown that Black women are less likely to be screened for mental health relative to White women (Sidebottom et al., 2021), have lower mental health care utilisation (Song et al., 2004), and are less likely to receive treatment (Kozhimannil et al., 2011). We observe similar racial disparities in our sample. 24% of non-Hispanic Black control mothers had a mental health diagnosis, in comparison to 43% of non-Hispanic White control mothers. 21% of White control mothers had some medication for mental health, relative to 10% of Black control mothers. We examine whether access to NFP can alleviate these racial differences in mental health outcomes. We do not find evidence to this effect. In fact, we find that Non-Hispanic Black mothers assigned to NFP were less likely to receive medication than their counterparts in the control group, by 3 percentage points (95% confidence interval: -5.02, -0.36). In

contrast, non-Hispanic White mothers assigned to NFP were more likely to receive medication than control mothers, by 4 percentage points (95% confidence interval: 0.17, 8.53). While racial subgroup analysis was pre-specified in the pre-analysis plan, these effects are not robust to controlling for the testing of multiple hypotheses.

Given the diverging mental health experiences of Black and White mothers, we explore potential drivers of this heterogeneity by comparing baseline characteristics, participation in the NFP program, and mental health screening and referrals by the health system across Black and White mothers. Black mothers were in fact more likely to present with depressive symptoms at intake (21% vs. 17%). While Black mothers had fewer NFP visits than White mothers, consistent with the prior literature (O'Brien et al., 2012), differences in engagement with the program are small and unlikely to drive differing experiences with the program.

In order to better understand how the program affects mental health diagnoses and treatments, we also consider how the program may change the likelihood of being screened for depression. The U.S. Preventive Services Task Force (USPSTF) recommends screening for depression for all adults, including pregnant and postpartum women, as well as having systems in place such that individuals can receive sufficient follow-up and treatment if needed (O'Connor et al., 2019). Specifically for mothers, the USPSTF recommends counseling intervention for mothers who are at increased risk of depression in the pregnancy or postpartum period, defined as the perinatal period (United States Preventive Services Task Force, 2019). We find that Black mothers participating in NFP were less likely to be screened at recommended intervals and referred to mental health services by NFP nurses.

We contribute to several strands of the literature. Evidence on the impacts of intensive nurse home visiting programs on diagnosis, treatment, and occurrence of mental health conditions, substance use, and domestic abuse is somewhat mixed. There is some evidence that nurse home visiting can have a positive effect on reducing depressive symptoms and use of psycholeptics (Sandner et al., 2018; Michalopoulos et al., 2019), but most studies have found no significant effects of nurse home visiting on maternal depression outcomes (Jacobs et al., 2015; Olds et al., 2004b). Few U.S. trials of the Nurse Family Partnership (NFP) have found any treatment effects on substance use

and domestic abuse (Olds et al., 2019, 2004a,b). Canadian and Dutch trials of NFP found reductions in prenatal smoking (Mejdoubi et al., 2014; Catherine et al., 2020), and physical assault rates (Mejdoubi et al., 2013). We advance this literature by providing new experimental evidence that combines survey data with rich administrative data on healthcare utilisation and treatment. Previous evaluations of NFP in the US have relied on survey data to examine maternal mental health, substance use and domestic abuse. Sandner et al. (2018) have similar administrative data but study a different program (Pro Kind) and context (Germany); the German healthcare system differs substantially to the U.S. By linking detailed NFP survey data with healthcare claims data we are the first in the U.S. setting to explore a range of health related outcomes that reflect mothers' actual healthcare use.

We are the first to study nurse home visiting program effects on maternal mental health utilisation outcomes by race, which allows us to examine whether intensive nurse home visiting programs address racial inequities in mental health. Previous research has documented substantial racial disparities in rates of perinatal mental health diagnosis and treatment (Kozhimannil et al., 2011; Sidebottom et al., 2021), and large racial inequalities in maternal health within income bands (Kennedy-Moulton et al., 2022). Research on how to address these disparities is lacking.

More broadly, we contribute to the literature on how public policies and welfare programs affect maternal mental health. A more generous social safety net has positive impacts on maternal mental health (Schmidt et al., 2023), as do tax credits (Evans and Garthwaite, 2014; Boyd-Swan et al., 2016; Lenhart, 2019; Gangopadhyaya et al., 2020), and parental leave (Persson and Rossin-Slater, 2024; Bütikofer et al., 2021). Additionally, Medicaid expansions for both the mother (Guldi and Hamersma, 2023) and child (Grossman et al., 2022) have been found to reduce maternal depression. Our paper extends this literature by evaluating the value on maternal mental health of a different type of benefit - having regular access to a healthcare professional in the home. NFP is time and resource intensive, and so our finding of null effects overall is important evidence for policymakers assessing the merits of different public policy options.

The paper proceeds as follows. Section 2 provides background on the Nurse Family Partnership. We discuss methods in Sections 3 and 4, then present our

empirical results in Section 5. In Section 6 we explore mechanisms. Section 7 provides a discussion and concludes.

## 2   Nurse-Family Partnership

NFP is an established home-visiting program for low-income, first-time mothers and their families. The original trials evaluating NFP took place between the 1970s and 1990s (Olds, 2006), with families enrolled in these trials being followed across many decades (Eckenrode et al., 2017; Conti et al., 2024; Olds et al., 2019). Evidence from these early trials motivated the scale-up of NFP to forty US states and eight countries with millions of dollars in U.S. federal funding through the Maternal, Infant, and Early Childhood Home Visiting (MIECHV) program.

As part of the NFP program, registered nurses visit families in their homes every other week during pregnancy and every month for up to two years following birth. Visits last between sixty to ninety minutes, and take place weekly for the first four weeks after enrolment and then fortnightly until the birth. This totals up to 40 home visits - up to 15 visits during the prenatal period, up to 8 visits in the postpartum period through to 60 days post delivery, and up to 9 visits in the remaining period until the child is 24 months. The visit cadence can be adjusted depending on mothers' needs, and families can choose to exit the program anytime. NFP nurses use methods such as motivational interviewing, goal-setting, and educational tools to provide the most value to the mothers they see. The nurses routinely assess physical and mental health, and can make referrals to health and social services. Mothers could view nurses as a substitute for some forms of routine healthcare, although previous results from the trial have indicated limited differences in routine prenatal or postpartum care across treatment and control groups (Gourevitch et al. (2023); Rokicki et al. (2024)). Nurses are mandated reporters of child abuse and neglect. The NFP nurses tailor their visits and visit content to the clients' needs and preferences, while also building in specific program requirements. For example, nurses must screen mothers for depression and anxiety at pre-specified intervals (36 weeks pregnant, 1-8 weeks postpartum, 4-6 months postpartum, and 12 months postpartum) and follow protocols for referral and care coordination for mothers who screen positive. Screening is also recommended at intake, or

within the first few visits. Additional screening is encouraged, as needed. The recommended screening tools are the Patient Health Questionnaire-9 (PHQ-9) for depression, and the Generalised Anxiety Disorder (GAD-7) form for anxiety. The results of these screens are recorded in the program's database, along with any referrals made. Nurses can make referrals for mental health services by calling a community provider and requesting services, or can support mothers in connecting to providers themselves. Screening provided by NFP nurses that suggests depression or anxiety is not considered a clinical diagnosis of depression, and is not a substitute for a complete mental health assessment and clinical judgement. In addition, NFP nurses are not gatekeepers of mental health diagnosis and treatment. Mothers can directly seek out diagnosis and treatment from healthcare providers. With regards to substance use, nurses attempt to identify mothers who are currently smoking, drinking alcohol, and/or using illicit substances given the risks that these behaviours pose for infant health, and encourage cessation of such activities. Nurses can also provide resources for cessation, such as information on hotlines, and refer mothers to healthcare providers. In addition, nurses have a unique opportunity to identify cases of intimate partner violence given that visits typically occur in the home. Furthermore, nurse are able to build a relationship with mothers throughout pregnancy and postpartum. Nurses should complete a clinical intimate partner violence (IPV) assessment during pregnancy, between the 5th and 7th visit. If the mother is diagnosed with IPV at this first assessment, or if the mother is engaging with a new partner, or re-engaging with a former abusive partner, their nurse should complete the assessment again at the 12 weeks postpartum visit. Nurses should repeat the assessment 13 to 18 months post birth. At any point, nurses can delay this assessment to ensure the mother's privacy and safety. If nurses identify a mother as being at risk, they can refer the mother to healthcare providers and locally available services. These services are likely to differ substantially across NFP sites as there aren't standardised interventions, resulting in varying support for mothers experiencing domestic violence.

The South Carolina NFP trial was delivered through the South Carolina Department of Health and Human Services (SCDHHS) with initial funding via a Medicaid 1915(b) waiver and philanthropists. The program was financed through a Pay for Success model, so that later funding was delivered

by SCDHHS if evaluation of the trial found impact estimates that exceeded pre-specified thresholds [1]. NFP was well-established in South Carolina, as it had operated in the state since 2009. This means that referral pathways were already embedded prior to the trial. The trial was implemented by nine agencies, each serving their own catchment area, within the state. Both rural and urban areas were covered by the program during the trial. Prior to the RCT, NFP served 500 to 600 mothers annually in South Carolina. The trial increased this number to 1200 mothers annually, with a respective increase in staffing.

Before analysis, maternal mental health was pre-specified as a key outcome variable. Other additional pre-specified outcomes for the South Carolina trial have been explored in other papers. These found that NFP had no significant effects on birth or neonatal outcomes (McConnell et al., 2022), prenatal care (Gourevitch et al., 2023), postpartum care (Rokicki et al., 2024) or birth spacing and the take-up of contraception (Steenland et al., 2024). Child health and social service utilisation outcomes are being explored in other forthcoming manuscripts.

## 3 Experimental design

### 3.1 Study design

We conducted an individually randomised clinical trial of NFP in South Carolina (Baicker et al., 2021). Trained NFP staff members ran trial enrolment. Potential study participants were identified by the South Carolina Department of Health and Human Services based on the Medicaid eligibility database. Others were recommended via referring partners (e.g. local heath care providers, schools, WIC agencies), or by a family member or friend. Individuals could also self-refer. Following referral, eligibility had to be determined. Individuals were eligible if they were currently pregnant (with gestation period of less than 28 weeks), had no previous live births, were income-eligible for Medicaid during pregnancy [2], were at least 15 years old, lived in the NFP catchment area, were not incarcerated or living in a lockdown facility, and had sufficient

---

[1] More detail on the Pay for Success model can be found in the study protocol (McConnell et al., 2020).

[2] Note that assignment to NFP does not affect Medicaid status.

language fluency to benefit from the program. Study participants then completed a baseline survey. Following this, study participants were randomised into the treatment group, with access to the NFP program, or the control group. Two-thirds of mothers were randomised to the treatment group, and a third to the control group. The control group had access to regular care available in South Carolina, including community and medical services. Mothers in the control group may have received services during pregnancy that mirror some of what NFP offers. For example, the Healthy Families America, Healthy Steps and Parents as Teachers programs. Most of these programs enrol after pregnancy, and operate at a much smaller scale than NFP during the study period. Further, mothers enrolled in Medicaid have access to two postpartum home visits up to six weeks after delivery. These home visits includes a medical assessment of the mother and infant, an assessment of barriers to health, counselling on postpartum recovery, family planning, information on the needs of a newborn, and assistance with establishing a primary source of care for the mother and infant. Only 11% of control group participants received at least one of these visits, while 6% of NFP intervention participants received at least one of these visits (Rokicki et al., 2024). All participants were given a list of available community resources for first time, low-income mothers. Most home visits took place in the client's home, but there were disruptions caused by the COVID-19 pandemic. South Carolina restricted public gatherings on March 23rd 2020, and following this date, 93% of home visits were via telehealth. 88% of mothers concluded their pregnancy prior to the start of the COVID-19 pandemic, and 84% of mothers concluded their 60 day postpartum period prior to the start of the pandemic. The nurses spoke English or Spanish, and translation services were available for mothers who spoke another language. The nurses in our trial had a mean caseload of 14.35 mothers per month, and a median monthly caseload of 12.72 mothers. In an original NFP trial, visitors had a caseload of 25 mothers (Olds et al., 2004b). More details on the trial are available in the published study protocol (McConnell et al., 2020).

## 3.2 Data

We combine several administrative datasets to evaluate the effects of NFP on maternal mental health, substance use, and domestic violence. First, we use data from a baseline survey that collected information on participants'

demographic and health information, such as age, race and ethnicity, body mass index, education, receipt of social service programs, substance use, and mental health symptoms and treatment history. Next, we use NFP program data. This includes information on visits conducted by the nurses, mental health screening, and referrals to mental health and substance use providers. The NFP program data includes nurse identifiers and characteristics, such that we can identify the race of mothers' nurses. Third, we use vital records data to identify live births and fetal deaths and for linkage to other datasets. Fourth, we use Medicaid claims and hospital discharge data, including pharmacy records, to observe diagnoses of depression, anxiety, substance use, and domestic violence. We also see receipt of treatment (medication or therapy). Finally, we use data from the South Carolina Department of Mental Health to observe therapy visits. We cannot observe any healthcare use outside of Medicaid and the South Carolina Department of Mental Health. Participants are matched to administrative data via a probabilistic match based on name, race, Social Security number, birth date, and Medicaid ID in collaboration with South Carolina's data repository.

## 3.3 Sample

The sample is made up of study participants with a live birth in the vital records data within 120 days of the expected delivery date reported on the baseline survey. We call these "index births". Supplemental Appendix Table 1 provides more detail. We include all mothers in our index birth sample, regardless of whether they match to all of the relevant data sources. As a robustness check, we repeat the program analysis on a sample of mothers who have at least one day of Medicaid coverage in the postpartum period and on a sample of mothers who have continuous Medicaid coverage up to 60 days postpartum.

There are no statistically significant differences in baseline characteristics between the treatment and control groups (Table 1). The sample is racially diverse: 55% of mothers self-identified as Black, 35% as White, and 10% as Hispanic. 18% of the sample are teenagers between the ages of 15 and 17, and just over half (54%) are 19 to 24 years old. Considering highest level of education, a little less than 60% of the sample have at most a high school diploma. Nearly 20% of the sample report depressive symptoms at baseline,

and 66% report high levels of stress. 14% of mothers report receiving mental health treatment in the past year. Half of mothers (51%) drank alcohol in the 3 months prior to their pregnancy, and a quarter (26%) of mothers smoked cigarettes in the same time period. 66% of the sample received at least one social service program such as Temporary Assistance for Needy Families or the Supplemental Nutrition Assistance program.

As we do subgroup analyses by race, we also check if there are statistically significant differences in baseline characteristics across the treatment and control arms for non-Hispanic White and Black mothers separately. Supplemental Appendix Table 7 show that there is balance across the arms for these two subgroups.

Table 1. Balance table

| Variables | Treatment Mean | Control Mean | Difference P-val |
|---|---|---|---|
| Age (%) | | | |
| - 15-18 | 18.21 | 17.53 | 0.56 |
| - 19-24 | 54.81 | 54.98 | 0.91 |
| - 25-34 | 24.16 | 25.47 | 0.31 |
| - 35+ | 2.82 | 2.02 | 0.09 |
| Race/ethnicity (%) | | | |
| - Hispanic | 5.52 | 6.25 | 0.32 |
| - Non-Hispanic Black | 55.04 | 55.62 | 0.71 |
| - Non-Hispanic White | 34.89 | 34.39 | 0.73 |
| - Non-Hispanic Other | 4.55 | 3.75 | 0.21 |
| BMI (%) | | | |
| - <18.5 | 5.73 | 5.34 | 0.58 |
| - 18.5-24.9 | 37.18 | 39.52 | 0.12 |
| - 25-29.9 | 22.22 | 21.66 | 0.66 |
| - >30.0 | 34.87 | 33.48 | 0.34 |
| Highest education level (%) | | | |
| - < HS | 22.62 | 21.70 | 0.47 |
| - HS | 36.00 | 34.27 | 0.23 |
| - < BA | 33.89 | 36.05 | 0.13 |
| - BA+ | 7.47 | 7.97 | 0.53 |
| Mental health/substance use (%) | | | |
| - Depressive symptoms | 19.27 | 18.70 | 0.63 |
| - High stress | 66.06 | 66.01 | 0.97 |
| - MH treatment | 13.79 | 13.54 | 0.81 |
| - Alcohol use | 50.72 | 51.32 | 0.69 |
| - Cigarette use | 25.05 | 27.03 | 0.14 |
| Other health indicators (%) | | | |
| - ED use | 51.11 | 52.26 | 0.44 |
| - Fair/poor health | 12.79 | 11.04 | 0.08 |
| Family planning indicators (%) | | | |
| - FP/birth control access | 55.88 | 55.42 | 0.76 |
| - Desire for children | 67.62 | 67.73 | 0.94 |
| Pregnancy indicators at baseline (%) | | | |
| - Gestational age | 1486.77 | 1452.05 | 0.07 |
| - Started prenatal care (%) | 85.08 | 84.04 | 0.34 |
| Socioeconomic indicators (%) | | | |
| - Used social services | 65.46 | 66.81 | 0.35 |
| - Worked for pay | 52.94 | 52.63 | 0.83 |
| - Lived with parents | 43.30 | 44.34 | 0.49 |
| - Housing insecure | 16.50 | 17.51 | 0.37 |
| Observations | 3295 | 1637 | 4932 |

Notes: Survey data collected after enrolment and before randomisation between 2016 and 2020. "Non-Hispanic Other" is Asian, Indigenous, Native Hawaiian/Pacific Islander and more than two races. BMI is body mass index. '<HS' is less than a high school diploma. "HS" is a high school diploma or equivalent. "<BA" is some college, less than a bachelor's degree. "BA+" is a bachelor's degree or higher. "Depressive symptoms" is a Patient Health Questionnaire 2 score of >= 3. "High stress" is a Perceived Stress Scale 4 score of >=4. "MH treatment" is receipt of mental health treatment in the last year. "Alcohol use" is reported alcohol use in the 3 months before pregnancy. "Cigarette use" is reported cigarette use in the 3 months before pregnancy. "ED use" is use of the emergency department in the 6 months before enrolment. "Fair/poor health" is self-reported health as fair or poor. "FP/birth control access" is whether the mother has access to family planning or birth control. "Desire for children" is self-reported desire for children in the future. "Gestational age" is at enrolment in weeks. "Started prenatal care" is receipt of at least one prenatal visit before enrolment. "Used social services" is use of programs such as TANF; SNAP; SSI; WIC; and unemployment benefits. "Housing insecure" is whether the mother moved at least twice in the previous 12 months.

### 3.4 Outcomes

In order to analyse the effects of NFP on perinatal mental health, we look at outcomes relating to diagnosis and treatment of mental health conditions in the perinatal period. We also examine mental health adjacent outcomes: substance use and maternal experience of violence. We focus on outcomes that were pre-registered in the clinical trials registry, and set out in the trial protocol and pre-analysis plan (McConnell et al., 2020, 2021). These pre-specified outcomes include: any outpatient diagnosis of depression, anxiety or stress-reaction; any medication; any psychotherapy visit; any inpatient or emergency department visit with a diagnosis for depression, anxiety, or stress-reaction; the number of inpatient or emergency department visits with a diagnosis for depression, anxiety, or stress reaction; whether the mother received a second prescription or therapy visit within 120 days of their first; any substance use (neonatal abstinence disorder, opioids, tobacco, alcohol, sedative, cocaine, amphetamines, cannabis), and any maternal experience of violence or homicide. A composite outcome captures any depression, anxiety, or stress-reaction diagnosis; any psychotherapy visit; or any medication. Any medication is defined as a filled prescription for antidepressants or anxiolytics. ICD-10 diagnosis, therapeutic class codes, and Current Procedure Terminology (CPT) codes used to define these outcomes can be found in Supplemental Appendix Tables 2, 3, 4, 5, 6. The diagnosis outcome is intended to capture detection of mental health conditions. The variables including prescriptions and therapy measure mental health treatment for mothers. The inpatient or emergency department visit outcome represents an outcome the program might hope to avoid, as these hospital visits could indicate inadequate outpatient treatment of perinatal depression. All outcomes are examined throughout the pregnancy period and up to sixty days postpartum, as this is when many women lost Medicaid coverage in South Carolina during the trial (prior to the onset of the COVID-19 pandemic). The variables measuring inpatient and emergency department visits with a diagnosis for depression or anxiety exclude inpatient claims for labour and delivery, as diagnoses at delivery may arise from depression screening at delivery or represent coding as a comorbidity or secondary diagnosis. These analytical choices were pre-specified in a pre-analysis plan (McConnell et al., 2021).

Using healthcare claims, discharge data and prescription fills is an imper-

fect way to assess prevalence of depression and anxiety in the study population. Receiving a depression diagnosis by a Medicaid provider is a function of the patient's presence at the facility, the information revealed to a doctor by the patient, and the doctor's propensity to diagnose to that patient. Further, seeing an increase in depression diagnoses could represent either an increase in the prevalence, or an increase in the identification of depression in a population. Participation in NFP could reduce depressive symptoms through easing stress and anxiety via interactions with NFP nurses, or it could increase the detection of perinatal depression if the NFP nurses helped mothers recognise their symptoms and encouraged them to seek care and treatment from a mental health provider. The former would lead to a reduction in depression diagnoses, and the latter, an increase. Moreover, medication use is observed for mothers who filled a prescription (although we cannot confirm whether the medication was taken as prescribed). We are not able to observe mothers who received a prescription but did not fill it. Finally, we cannot observe treatment that is not billed to Medicaid. Therefore, the outcomes we explore can be thought of as indicating health care utilisation relating to mental health rather than indicators of mental health status.

In order to understand the context of routine depression, substance use, and domestic violence screening within the Medicaid program in South Carolina, we also conduct descriptive analysis on exploratory outcomes. These exploratory outcomes include mental health screening and referrals captured within Medicaid claims, and indicators of program implementation related to what happened during NFP visits. We look at perinatal mental health screening through routine health care delivery in the Medicaid program. We observe screening for mental health, alcohol and substance use, smoking patterns, and domestic violence using the Screening, Brief Intervention and Referral to Treatment (SBIRT). Clinicians are encouraged to deliver SBIRT to all pregnant mothers on Medicaid, as early as possible. The first component of the SBIRT is screening, where a doctor asks a set of questions to assess the risks to the mother and infant. This takes five to ten minutes to complete. If needed, a brief intervention is then conducted. This is a short counseling session to highlight potential issues and motivate the mother to acknowledge any problems. The brief intervention can last between five to thirty minutes. Then, there is referral to treatment. As per SBIRT guidelines, if an individual screens pos-

itive for any risk factors, providers should make a referral and ideally secure an appointment before the individual leaves. Patients can be referred to the Department for Mental Health, Department of Alcohol and Other Drug Abuse Services, Quitline (for tobacco cessation), a domestic violence hot line, or a private provider (South Carolina Department of Health and Human Services, 2010). Providers can bill for the screening once per fiscal year, and the brief intervention twice per fiscal year. We also observe specific depression screening in Medicaid claims. NFP recommends using a standardised depression screening tool for postpartum women - the Edinburgh Postnatal Depression Scale (EPDS). The EPDS is a screening questionnaire that aims to identify mothers with symptoms of depression and anxiety in the perinatal period. The EPDS is not a diagnostic tool; mothers with scores indicating depressive symptoms should receive follow-up care. Mothers complete the questionnaire themselves, answering ten questions about their emotional experiences over the past week. The American College of Obstetricians and Gynecologists (ACOG) recommends that all women are screened for perinatal mental health conditions, at the following cadence: at the first obstetric visit to assess mental health before pregnancy, at 24 to 28 weeks of gestation to assess mental health during pregnancy, and at the comprehensive postpartum visit to identify onset in early postpartum (The American College of Obstetricians and Gynecologists, 2024). Screenings can take place during prenatal care appointments. Providers can also bill for a depression screening at well-child visits under South Carolina's Medicaid policy (twice per date of service) (South Carolina Department of Health and Human Services, 2024). Infants should receive a well-child visit 3-5 days, one month, and two months after birth.

In addition to screening in routine and emergency health care utilisation, mothers in the treatment group could also be screened and referred for services by NFP nurses. We look at the prevalence of depression screening, mental health referrals, and substance use referrals by NFP nurses across the pregnancy and postpartum period in the treatment arm. We examine outcomes in the prenatal period, from birth to sixty days postpartum, and a combination of these two periods.

Finally, to examine whether the NFP program reduces racial disparities, we look at screening and referral within the Medicaid program and screening accordance to timing guidelines by nurses implementing NFP, by race. For

14

the study period we focus on, NFP guidance advised screening at intake, at 36 weeks pregnant, and 1-8 weeks after birth.

## 3.5 Descriptive analysis of program implementation

Table 2 shows descriptive statistics on participation in the South Carolina NFP trial across the prenatal period and up to 60 days postpartum. 98% of mothers received at least one nurse visit in the prenatal period, and this figure is 76% in the postpartum period. Telehealth visits were less common - around a quarter of mothers received at least one telehealth visit. Visits lasted about an hour, and in-person visits were over double the length of telehealth visits. The mean number of visits in the prenatal period was 9.34, and 3.64 in the postpartum period up to 60 days after birth.

Table 2. Participation in NFP

|  | Prenatal %/Num. | 60 days postpartum %/Num. |
| --- | --- | --- |
| **program encounters** | | |
| Received at least one visit | 98.27 | 75.51 |
| Received at least one in-person visit | 98.24 | 66.43 |
| Received at least one telehealth visit | 28.01 | 25.25 |
| Received visits to within 14 days of delivery | 78.26 | 70.83 |
| Number of visits | 9.34 | 3.64 |
| Number of in-person visits | 8.29 | 2.89 |
| Number of telehealth visits | 1.05 | 0.75 |
| **Visit duration** | | |
| Avg. duration of visits | 66.82 | 62.97 |
| Avg. duration of in-person visits | 70.46 | 70.10 |
| Avg. duration of telehealth visits | 28.16 | 28.91 |

Notes: N=3,295. Data only available for participants who matched to program data. "Duration" variables available for participants with at least one in-person visit. "Number" and "Duration" variables are means. "Received" variables are percentages. "Duration" variables are in minutes.

## 3.6 Descriptive analysis of mental health screening and referral

Table 3 Panel A shows depression screening in the NFP program. Nearly all (99%) mothers in the treatment group received a depression screening at some point from the start of their pregnancy up to sixty days postpartum. Among mothers in the treatment group, 16% were ever referred to treatment, with the majority of referrals occurring in the prenatal period. A small proportion of NFP mothers received a referral for substance use, 4%, with the majority of these referrals occurring in the prenatal period when there is heightened

concern for the health of the infant from maternal substance use. Table 3 Panel B shows screening by the Medicaid program for the full sample. The rate of depression screening by Medicaid providers is much lower than in the NFP program. There is almost no depression screening by Medicaid providers during pregnancy. After birth, 24% of mothers receive a depression screening. This figure is low given the guidance to screen in the postpartum period (The American College of Obstetricians and Gynecologists, 2024), and the fact that providers can bill depression screens for mothers during well-child visits. Providers are using the SBIRT screen in the prenatal period - 30% of mothers received a prenatal SBIRT screening and 8% of mothers receive an SBIRT referral. After delivery, the use of SBIRT screens falls close to zero.

Table 3. Prevalence of mental health screening and referral

|  | Prenatal | 60 days pp |
|---|---|---|
| **Panel A: NFP** | | |
| Any depression screening | 97.32 (16.14) | 84.02 (36.65) |
| Any referral for mental health | 12.18 (32.71) | 6.86 (25.27) |
| Any referral for substance use | 3.57 (18.55) | 0.51 (7.15) |
| Any referral for domestic violence | 1.20 (10.90) | 0.30 (5.47) |
| **Panel B: Medicaid** | | |
| Any depression screening | 0.53 (7.24) | 23.95 (42.68) |
| Any SBIRT screening | 29.89 (45.78) | 0.28 (5.32) |
| Any depression or SBIRT screening | 30.37 (45.99) | 24.15 (42.80) |
| Any SBIRT referral | 7.91 (26.99) | 0.12 (3.49) |

Notes: N=2,509 for "Panel A: NFP - Prenatal" sample. N=2,315 for "Panel A: NFP - 60 days pp" sample. N=4,932 for "Panel B: Medicaid - Prenatal/60 days pp" sample. "Panel A: NFP" denotes screening and referral conducted by NFP nurses to mothers in the treatment group. "Panel B: Medicaid" denotes screenings and referrals that are observed in Medicaid claims data, for both treatment and control mothers. For all variables the values represent percentages. "Prenatal" denotes that the screening or referral took place during the prenatal period. For "Prenatal", Panel A statistics are calculated for the sample of mothers who had NFP visits until birth. "60 days pp" denotes that the screening or referral took place after delivery, up to 60 days postpartum. For "60 days pp", Panel A statistics are calculated for the sample of mothers who had NFP visits until 60 days postpartum. "NFP - Any referral for domestic violence" is whether the mother received a referral for maternal experience of violence or homicide. "Medicaid - Any depression screening" is whether the mother received a depression screen (Edinburgh Postnatal Depression Scale) during any well-child visit. "Medicaid - Any SBIRT screening" is whether the mother received the "Screening, Brief Intervention, and Referral to Treatment" screen. "Medicaid - Any SBIRT referral" is whether the mother was referred to a specialist from an SBIRT screen. Standard errors in parentheses.

# 4 Empirical approach

To evaluate the effects of NFP on maternal mental health, we use an intent to treat (ITT) framework that estimates the average impact of NFP comparing outcomes for all index births randomly assigned to the NFP treatment group and all index births randomly assigned to the control group.

We wish to estimate the causal effect of access to NFP on mothers' mental health outcomes. To do so, we estimate the follow equation:

$$Y_i = \alpha + \beta \cdot NFP_i + \gamma \cdot X_i + \eta \cdot M_i + \epsilon_i \tag{1}$$

where $i$ is an individual mother. $Y_i$ denotes a perinatal mental health outcome described in Section 3.4. $NFP_i$ is a binary indicator for whether the mother is randomly assigned to the NFP treatment group. We control for the following pre-specified individual level characteristics $X_i$: indicators for the implementing NFP agency, age groups, race and ethnicity, gestational age at time of study enrolment, relationship with father of the child, education, employment, receipt of social services, housing stability, health status, psychological state, access and utilisation of health care, alcohol and smoking use, stress, pregnancy risk factors, and family planning metrics. We account for missing baseline covariates using a dummy-variable adjustment approach (Puma et al., 2009). Specifically, we set missing covariate values to a constant value and add dummy variables $M_i$ for missing values. We estimate this equation using ordinary least squares and estimate robust standard errors. We also present estimates that omit covariates and missing covariate dummy variables.

Our primary specification relies on an ITT framework rather than a treatment on the treated (TOT) framework because the policy relevant analysis seeks to understand the impact of access to the NFP program. Many families in a realistic implementation setting may choose not to participate in the full program activities. We see almost total take-up of at least one NFP visit (see Table 2). Beyond this, the variation in the level of program participation is what we would expect in any realistic implementation. In Supplemental Appendix Table 12, we estimate the TOT effects by instrumenting whether the mother received NFP visits to 60 days postpartum with randomisation into the NFP treatment arm.

We conduct subgroup analyses for non-Hispanic Black mothers, non-Hispanic White mothers, and a vulnerable subgroup. These subgroups were specified in the pre-analysis plan. The vulnerable subgroup is defined as participants who were at least one of the following: younger than 19 years at baseline, without a high school degree, or who experienced mental health challenges. Mothers with mental health challenges were identified by either receiving a score of three or higher on the Patient Health Questionnaire 2, or reporting receipt of

mental health treatment in the year prior to enrolment. These indicators of vulnerability put the mother at risk of developing postnatal depression (Biaggi et al., 2016; Goyal et al., 2010; Kingston et al., 2012). Additionally, this subgroup mimics subgroups in which the original NFP trials found greater impacts. Subgroup analyses by race were pre-specified as prior work shows racial differences in treatment for perinatal depression (Kozhimannil et al., 2011).

## 4.1 Threats to validity

### 4.1.1 Internal validity

Our treatment group is defined as mothers randomised to access to NFP. All treatment mothers will thus be "treated" and the compliance assumption satisfied. However, we cannot ensure that all treated mothers remain in NFP. We can ensure no defiers. This means that there will be some degree of variation in exposure to NFP nurse visits within the treatment group. The treatment-on-the-treated estimates in Supplemental Appendix Table 12 show the impact of NFP among those who attended NFP visits to 60 days postpartum.

A related point discussed in Section 2 is that some control mothers may have received community services that were similar to those offered by the NFP program, although the scale and intensity of these benefits were likely small. This may limit our ability to detect the full impact of NFP services.

One may be concerned about spillover effects, say, if mothers in NFP can transmit information they have learned from their NFP nurse to control group mothers. However, we consider that any possible spillovers would not have meaningful effects on maternal outcomes, as the NFP program is tailored to individuals and thus the nurses' advice and support is mother specific.

Finally, as mothers in the control group were aware that they did not receive the program, this may have altered their behaviour.

### 4.1.2 External validity

Next, we consider the external validity of our results to other populations. Specifically, given that NFP is targeted at low-income mothers, populations that are likely to receive NFP. For example, Medicaid eligible families.

Our NFP trial sample differs from the broader female Medicaid population. Our sample had a larger share of Black mothers (55%) than the general female

Medicaid population (25%), and a lower share of Hispanic mothers (6% in our sample vs. 22% in all female Medicaid participants). A smaller proportion of our mothers report having less than a high school degree than Medicaid women more broadly (22% vs. 35%) (Kaiser Family Foundation, 2022). These differences in characteristics suggest caution in extrapolating our results to a national Medicaid eligible population, particularly results by race.

Previous and current NFP trials typically serve a smaller group of families and each NFP trial has its own features. Programs are tailored to the local population for maximum impact. Thus expanding NFP access to larger groups of families will likely require harmonisation of program design and implementation across sites to maintain effectiveness at scale. One potential concern with expanding NFP access to more families is that potential participants may not be comfortable being research subjects and so would not enrol in a trial.

Finally, the U.S. context and the specific state in which the trial is run will affect the success and outcomes of the program, as healthcare provision and access varies by country and state. Thus we advise against generalising our results to a substantially different context.

## 5 Results

When we look at the pre-registered study outcomes, we find that the NFP program had no statistically significant effect on mental health diagnoses, nor on treatment in the form of medication or therapy. We also find no significant difference in substance use or experiences of maternal violence detected in a healthcare encounter. Table 4 shows the NFP program effects on our outcomes for all index birth mothers. Supplemental Appendix Table 8 shows the q-values calculated using the Benjamini and Hochberg (1995) procedure that accounts for multiple hypothesis testing.

Next, we examine heterogeneity by race to evaluate whether NFP can help reduce racial disparities in mental health outcomes. First, we document these racial disparities. In our sample, 43% of non-Hispanic White control group mothers received an outpatient mental health diagnosis in comparison to 24% of non-Hispanic Black control group mothers. Further, 21% of non-Hispanic White control group mothers received any medication, in comparison to 10% of non-Hispanic Black control group mothers. These racial disparities in rates of

medication use are similar to those documented in other studies; Huybrechts et al. (2013) found that 3.5% of Black mothers use antidepressants during pregnancy versus 14.4% of White mothers. We also find racial differences in therapy use: 6% of non-Hispanic White control group mothers received some therapy, relative to 4% of non-Hispanic Black control group mothers. In addition, non-Hispanic White mothers were more likely to have follow up treatment.

We find suggestive evidence of heterogenous NFP program effects by race; Table 5 shows NFP program effects for the sub-samples of non-Hispanic White and non-Hispanic Black index birth mothers. Non-Hispanic White mothers randomly assigned to NFP were 4% (95% confidence interval: 0.17, 8.53) more likely to receive mental health medication treatment than those randomly assigned to the control group. In contrast, non-Hispanic Black mothers were 3% (95% confidence interval: -5.02, -0.36) less likely to receive medication if they were in the NFP treatment group rather than the control group. Both of these effects are statistically significant, and the difference between the adjusted treatment effects for non-Hispanic White and Black mothers is also statistically significant. However, the results are not statistically significant after accounting for multiple hypothesis testing (Supplemental Appendix Table 9 shows the q-values calculated using the Benjamini and Hochberg (1995) procedure.) We do not see heterogeneous treatment effects on diagnoses.

We find no significant effects of NFP on maternal mental health for the vulnerable subgroup (Supplemental Appendix Table 10). Sensitivity analysis on a sample of mothers with any Medicaid coverage in the postpartum period and on a sample of mothers with continuous Medicaid coverage up to 60 days postpartum can be found in Supplemental Appendix Tables 11 and 12. In Supplemental Appendix Table 13, we estimate the treatment-on-the-treated effects by instrumenting whether the mother received NFP visits to 60 days postpartum with randomisation into the NFP treatment arm. We find no statistically significant TOT effects either.

Table 4. Treatment effects of NFP, full sample

|  | Treatment Mean | Control Mean | Unadjusted ITT | ITT | |
|---|---|---|---|---|---|
| Any OP diagnosis/medication/therapy | 31.81 | 33.17 | -1.36 | -1.34 | [-3.80, 1.11] |
| Any OP diagnosis | 28.44 | 29.99 | -1.56 | -1.55 | [-3.95, 0.84] |
| Any Medication | 13.20 | 13.19 | 0.01 | 0.06 | [-1.84, 1.95] |
| Any Therapy | 5.01 | 4.64 | 0.36 | 0.41 | [-0.77, 1.59] |
| Any IP/ED visit | 13.93 | 12.34 | 1.59 | 1.47 | [-0.37, 3.31] |
| Num. IP/ED visits | 0.24 | 0.21 | 0.03 | 0.03 | [-0.02, 0.08] |
| Treatment follow up | 8.86 | 8.67 | 0.19 | 0.14 | [-1.39, 1.68] |
| Any substance use | 21.49 | 22.42 | -0.93 | -0.47 | [-2.57, 1.62] |
| Any domestic violence | 1.67 | 1.65 | 0.02 | 0.01 | [-0.75, 0.77] |

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Full sample N =4,932. "ITT" is the intent-to-treat effect. For the "Any" variables, the means and effects are percentages. For the "Num" variable, the means and effects are numbers. "Any OP diagnosis/medication/therapy" is any outpatient diagnosis for depression or anxiety, medication, or therapy. "Any IP/ED visit" is an inpatient or emergency department visit with a postpartum depression or anxiety diagnosis, excluding labour and delivery. "Num. IP/ED" is defined the same, but counts the number of visits. "Treatment follow up" is a second prescription or therapy visit within 120 days of first prescription or therapy visit. 95% confidence intervals in brackets. All "ITT" models control for implementing NFP agency, age groups, race and ethnicity, gestational age at time of study enrolment, relationship with father of the child, education, employment, receipt of social services, housing stability, health status, psychological state, access and utilisation of health care, alcohol and smoking use, stress, pregnancy risk factors, family planning metrics, and dummies for missing baseline covariates.

Table 5. Treatment effects of NFP, by race

| | White | | | Black | | | Difference |
|---|---|---|---|---|---|---|---|
| | Control Mean | ITT (1) | | Control Mean | ITT (2) | | (1)-(2) P |
| Any OP diagnosis/medication/therapy | 47.23 | 0.71 | [-4.02, 5.44] | 26.95 | -1.76 | [-5.10, 1.57] | 0.40 |
| Any OP diagnosis | 43.21 | -1.83 | [-6.55, 2.88] | 23.64 | -0.58 | [-3.78, 2.61] | 0.66 |
| Any Medication | 20.65 | 4.35 * | [0.17, 8.53] | 10.28 | -2.69 * | [-5.02, -0.36] | 0.00 |
| Any Therapy | 6.12 | 1.89 | [-0.53, 4.30] | 3.78 | -0.22 | [-1.70, 1.26] | 0.14 |
| Any IP/ED visit | 16.25 | 3.16 + | [-0.57, 6.88] | 10.99 | 0.56 | [-1.87, 3.00] | 0.25 |
| Num. IP/ED visits | 0.29 | 0.05 | [-0.06, 0.16] | 0.18 | 0.02 | [-0.03, 0.08] | 0.71 |
| Treatment follow up | 14.72 | 1.74 | [-1.77, 5.25] | 5.67 | -0.43 | [-2.19, 1.34] | 0.27 |
| Any substance use | 31.55 | -1.26 | [-5.31, 2.80] | 18.32 | 0.10 | [-2.72, 2.91] | 0.59 |
| Any domestic violence | 0.57 | 0.34 | [-0.52, 1.19] | 2.13 | -0.03 | [-1.27, 1.21] | 0.63 |

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

Notes: Non-Hispanic White sample N= 1,604. Non-Hispanic Black sample N= 2,551. "ITT" is the intent to treat effect. For all variables, the means and ITT are percentages. For the "Any" variables, the means and effects are percentages. For the "Num" variable, the means and effects are numbers. "Any OP diagnosis/medication/therapy" is any outpatient diagnosis for depression or anxiety, medication, or therapy. "Any IP/ED visit" is an inpatient or emergency department visit with a depression or anxiety diagnosis, excluding labour and delivery. "Num. IP/ED" is defined the same, but counts the number of visits. "Treatment follow up" is a second prescription or therapy visit within 120 days of first prescription or therapy visit. 95% confidence intervals in brackets. All models control for implementing NFP agency, age groups, race and ethnicity, gestational age at time of study enrolment, relationship with father of the child, education, employment, receipt of social services, housing stability, health status, psychological state, access and utilisation of health care, alcohol and smoking use, stress, pregnancy risk factors, family planning metrics, and dummies for missing baseline covariates.

# 6 Exploring mechanisms for differential programme experiences by race

Next, we investigate potential drivers of racial differences in mental health diagnoses and treatment, and program effects, by analysing the baseline survey, NFP program and medical claims data.

## 6.1 Racial differences in mental health prior to the intervention

Disparities in mental health treatment for Black and White mothers and a lower diagnosis rate for Black mothers could arise from racial differences in mental health status before the NFP intervention. Using data from the baseline survey, we investigate mental health symptoms upon intake, and treatment history by race. Non-Hispanic Black mothers were more likely than non-Hispanic White mothers to report depressive symptoms at intake during pregnancy, as exhibited in Table 6. 21% of non-Hispanic Black mothers scored at least three on the Patient Health Questionnaire 2, in comparison to 17% of non-Hispanic White mothers. Despite this difference in depressive symptoms by race, a much lower proportion of Black mothers had received mental health treatment in the year prior to NFP (8% relative to 24% of White mothers). Both of these racial differences are statistically significant. These findings hold when controlling for the nine different NFP Implementing Agencies (IA) and personal characteristics, as shown in Supplemental Appendix Table 14. Potential reductions in rates of mental health treatment for Black mothers participating in NFP do not seem to be driven by Black study participants experiencing fewer symptoms or being less in need of mental health treatment at baseline.

Table 6. Mental health status prior to intervention, by race

|  | Black (1) | White (2) | (1)-(2) |
|---|---|---|---|
|  | % | % | P |
| Depressive symptoms | 20.94 (40.70) | 16.74 (37.35) | 0.01 |
| High stress | 65.69 (47.49) | 65.49 (47.56) | 0.92 |
| Received mental health treatment in last year | 8.04 (27.19) | 24.24 (42.87) | 0.00 |

Notes: N= 1,705 for non-Hispanic Black mothers. N= 1,081 for non-Hispanic White mothers. Survey data collected after enrolment and before randomisation between 2016 and 2020. "Depressive Symptoms" is a Patient Health Questionnaire 2 score of at least 3. "High Stress" is a Perceived Stress Scale 4 score of at least 4.

## 6.2 Racial differences in participating in NFP

Differential NFP program effects by race could be driven by differing levels of participation in the NFP program. We examine this question in Table 7 Panel A. Non-Hispanic Black mothers were more likely to remain in NFP until birth, and there is no statistically significant difference in program retention when we consider the full study period from pregnancy to 60 days postpartum. Controlling for the NFP implementing agencies (IAs) and a rich set of personal characteristics confirms that there is no statistically significant difference in retention in NFP by race (Supplemental Appendix Table 15).

Another potential reason for the differential impact of the NFP program across race is because Black mothers received fewer, or shorter, nurse home visits. We explore whether there are racial differences in visit frequency and length. Table 7 Panel B shows that there is no difference in the average NFP home visit duration for non-Hispanic Black and White mothers. When we control for IA and personal characteristics (Supplemental Appendix Table 16), there is a small statistically significant reduction in the visit length of Black mothers relative to White mothers of 2 minutes. During pregnancy, the postpartum period up 60 days following birth, and the combined study period, Black mothers had statistically significantly fewer visits than their White counterparts. Across the combined study period, Black mothers have 1.21 fewer visits than White mothers. This effect remains in the postpartum period when adding in controls, as shown in Supplemental Appendix Table 16. Differences in the number of visits could reflect patient preferences, nurse preferences, or a combination of both. Differences could also reflect the nature of the relationship between the nurse and client.

## 6.3 Racial differences in mental health screening and referral patterns by NFP

Lower rates of mental health diagnoses and treatment amongst Black women could be due to differences in screening and referral for Black participants. Table 7 Panel C shows screening and referral outcomes for non-Hispanic Black and White mothers in the intervention group. Nearly all mothers were screened for depression by NFP nurses at least once across the study period. Although the absolute differences are small, Black mothers have statistically significantly

fewer depression screenings across the prenatal period and up to 60 days postpartum. According to program guidelines, NFP nurses should screen mothers at intake, at 36 weeks of pregnancy, and between 1 to 8 weeks following birth. Around 87% of mothers were screened at intake, with no significant difference by race. At 36 weeks into pregnancy ("Prenatal"), non-Hispanic White mothers were 9 percentage points more likely to be screened by NFP nurses than non-Hispanic Black mothers. At 1 to 8 weeks after birth ("60 days pp"), non-Hispanic White mothers were 4 percentage points more likely to be screened by NFP nurses than their Black counterparts. White mothers were also referred for mental health services at a higher rate than Black mothers. Supplemental Appendix Table 17 adds in controls for the NFP implementing agency and for a rich set of individual characteristics. Differences in mental health referrals between Black and White mothers remain.

## 6.4 Racial differences in mental health screening and referral patterns by Medicaid providers

Next, we consider whether lower rates of screening and referral for Black compared to White mothers observed in the NFP program data are also observed in screening and referral patterns in the Medicaid program. We revisit the prevalence of mental health screening during pregnancy and the postpartum period (as in Table 3) and break down the Medicaid analysis by race in Table 7 Panel D. We see that across the whole study period, there is no statistically significant difference in receipt of depression or SBIRT screening by race. Non-Hispanic White mothers, however, were significantly more likely to receive an SBIRT referral than non-Hispanic Black mothers. These differences cannot be explained by racial differences in Medicaid coverage, as shown in Supplemental Appendix Table 19. The dynamics that drive lower rates of mental health screenings in NFP do not appear to be present when screening is undertaken by Medicaid providers. This is consistent with limited differences in routine utilisation of prenatal and postpartum care reported in other manuscripts on this trial population (Gourevitch et al., 2023). Furthermore, we do not find statistically significant NFP treatment effects on Medicaid screening and referral outcomes for non-Hispanic Black and White mothers (see Supplemental Appendix Table 18).

Table 7. Visits, mental health screening, and referral in the NFP program and the healthcare system, by race

| | Prenatal | | | 60 days pp | | | Study period | | |
|---|---|---|---|---|---|---|---|---|---|
| | Black (1) | White (2) | (1)-(2) P | Black (1) | White (2) | (1)-(2) P | Black (1) | White (2) | (1)-(2) P |
| **NFP** | | | | | | | | | |
| *Panel A: Retention* | | | | | | | | | |
| Visits across period | 79.34 (40.50) | 75.76 (42.87) | 0.03 | 70.54 (45.60) | 69.94 (45.88) | 0.73 | 70.54 (45.60) | 69.94 (45.88) | 0.73 |
| *Panel B: Visits* | | | | | | | | | |
| Avg. visit length | 65.96 (22.52) | 66.23 (21.05) | 0.78 | 62.11 (36.59) | 61.17 (24.78) | 0.53 | 64.90 (22.23) | 64.39 (17.99) | 0.60 |
| Visit count | 11.05 (4.36) | 11.98 (4.51) | 0.00 | 5.09 (2.29) | 5.44 (2.33) | 0.00 | 16.35 (5.78) | 17.56 (5.97) | 0.00 |
| *Panel C: Screening and Referral* | | | | | | | | | |
| Number of depression screenings | 1.71 (0.67) | 1.82 (0.60) | 0.00 | 1.05 (0.79) | 1.14 (0.92) | 0.03 | 2.79 (1.11) | 2.98 (1.17) | 0.00 |
| Any depression screening | 96.67 (17.94) | 98.05 (13.85) | 0.06 | 82.78 (37.77) | 84.79 (35.94) | 0.24 | 98.75 (11.11) | 99.74 (5.14) | 0.02 |
| Screening adherence by nurse | 72.49 (44.68) | 81.81 (38.60) | 0.00 | 88.35 (32.09) | 92.20 (26.84) | 0.01 | 94.34 (23.11) | 96.16 (19.22) | 0.07 |
| Any mental health referral | 10.36 (30.48) | 15.02 (35.75) | 0.00 | 5.49 (22.79) | 8.99 (28.63) | 0.00 | 13.89 (34.60) | 19.84 (39.91) | 0.00 |
| **Medicaid** | | | | | | | | | |
| *Panel D: Screening and Referral* | | | | | | | | | |
| Any depression screening | 0.35 (0.06) | 0.94 (0.10) | 0.02 | 24.30 (0.43) | 22.13 (0.42) | 0.11 | 24.58 (0.43) | 22.94 (0.42) | 0.23 |
| Any SBIRT screening | 30.81 (0.46) | 30.42 (0.46) | 0.79 | 0.35 (0.06) | 0.25 (0.05) | 0.56 | 31.09 (0.46) | 30.67 (0.46) | 0.78 |
| Any depression or SBIRT screening | 31.13 (0.46) | 31.36 (0.46) | 0.87 | 24.58 (0.43) | 22.32 (0.42) | 0.10 | 48.37 (0.50) | 46.57 (0.50) | 0.26 |
| Any SBIRT referral | 6.82 (0.25) | 11.16 (0.31) | 0.00 | 0.04 (0.02) | 0.19 (0.04) | 0.13 | 6.82 (0.25) | 11.16 (0.31) | 0.00 |

Notes: N= 1,352 for non-Hispanic Black mothers who completed NFP visits to birth, sample for "NFP - Prenatal". N= 819 for non-Hispanic White mothers who completed NFP visits to birth, sample for "NFP - Prenatal". N=1,202 for non-Hispanic Black mothers who completed NFP visits to 60 days postpartum, sample for "NFP - 60 days pp"/"Study period". N=756 for non-Hispanic White mothers who completed NFP visits to 60 days postpartum, sample for "NFP - 60 days pp"/"Study period". N=2,551 for "Medicaid" sample of non-Hispanic Black mothers. N=1,604 for "Medicaid" sample of non-Hispanic White mothers. "NFP" denotes visits, screening, and referrals that are observed in NFP survey data, for treatment mothers. NFP survey data collected after enrolment and before randomisation between 2016 and 2020. "Medicaid" denotes screenings and referrals that are observed in Medicaid claims data, for both treatment and control mothers. "Prenatal" denotes that the outcome took place during the prenatal period. For the NFP variables, the "Prenatal" statistics are calculated for the sample of mothers who had NFP visits until birth. For the Medicaid variables, the "Prenatal" statistics are calculated for all mothers. "60 days pp" denotes that the outcome took place after delivery, up to 60 days postpartum. For the NFP variables, the "60 days pp" statistics are calculated for the sample of mothers who had NFP visits until 60 days postpartum. For the Medicaid variables, the "60 days pp" statistics are calculated for all mothers. "Study period" denotes the prenatal period and up to 60 days postpartum. For the NFP variables, the "Study period" statistics are calculated for the sample of mothers who had NFP visits until 60 days postpartum. For the Medicaid variables, the "Study period" statistics are calculated for all mothers. The statistics for "Visits across period", "Any depression screening", "Screening adherence by nurse", "Any mental health referral", 'Any SBIRT screening", "Any depression or SBIRT screening", and "Any SBIRT referral" are percentages. "Visits across period' is whether the mother received NFP visits within 14 days of the index birth for the "Prenatal" columns, and whether the mother received NFP visits up to 60 days postpartum for the "60 days pp" and "Study period" columns. "Avg. visit length" is the average (mean) NFP visit length in minutes for completed visits. "Visit count" is the number of at home NFP completed visits. "Screening adherence" is whether the mother was screened at 36 weeks pregnant for "Prenatal", whether the mother was screened 1 to 8 weeks after birth for "60 days pp", and whether the mother was screened at 36 weeks pregnant or 1 to 8 weeks after birth for "Study period". "Any depression screening" is whether the mother received a depression screen (Edinburgh Postnatal Depression Scale) during any well-child visit. "Any SBIRT screening" is whether the mother received the "Screening, Brief Intervention, and Referral to Treatment" screen. "Any SBIRT referral" is whether the mother was referred to a specialist from an SBIRT screen. Standard errors in parentheses.

## 6.5 Racial concordance between the nurse and mother

Finally, we conduct descriptive analyses to examine whether racial concordance between mother and nurse is associated with different NFP program outcomes for Black and White mothers. Prior work has found that racial concordance can result in increased take-up of preventive care and improved outcomes for Black patients (Alsan et al., 2019; Greenwood et al., 2018, 2020; Frakes and Gruber, 2022; Hill et al., 2023).

The majority of nurses are White (34%), 15% are Black, and the remainder are of unknown race. We define the mother's "main" nurse as the nurse that they saw the most frequently. 68% of treated mothers only saw one nurse during the study period. The median was 1 nurse, and the mean was 1.36. These figures are consistent across the full sample and non-Hispanic Black and White subgroups. Table 8 shows the extent to which Black and White mothers had racially concordant and discordant nurses.

The racial concordance analysis is presented in Table 9. The estimated effect of having a Black nurse instead of a White nurse for White patients is given by the coefficient on "Black Nurse" ($\beta_2$), and for Black patients it is given by the sum of the coefficients on "Black Nurse" and the interaction term ($\beta_2 + \beta_3$). The coefficient on the interaction term ($\beta_3$) denotes the differential impact of nurse race on outcomes of Black mothers relative to White mothers. We find evidence of increased participation in NFP, longer visits, and more depression screening for Black mothers with Black nurses relative to White mothers with White nurses. We find no statistically significant differences in visit length and mental health referrals. Black mothers who see Black nurses have 3.93 minute shorter visits than Black mothers who see White nurses. They also receive more depression screenings than Black mothers who see White nurses. The magnitude of this effect is small (0.22 screens) but represents an 9% rise relative to the mean. Black mothers who see Black nurses are also 5 percentage points less likely than Black mothers who see White nurses to receive a mental health referral, a 31% decline relative to the mean. Amongst Black mothers, our results suggest that seeing a Black nurse is not associated with significant improvements in program adherence, and a substantial reduction in the probability of receiving a mental health referral despite a slightly higher chance of being screened. Nurses are not randomly assigned to mothers and mothers can switch nurses, so we may be concerned

about selection into nurse race. To address concerns about causality, we add nurse fixed effects; the coefficient estimates are robust to their inclusion.

Table 8. Racial composition of nurses and mothers

|  | Full sample % | Black % | White % |
|---|---|---|---|
| **Main nurse** | | | |
| Black | 18.02 (38.44) | 21.41 (41.03) | 14.39 (35.12) |
| White | 53.43 (49.89) | 47.53 (49.95) | 60.89 (48.82) |
| Race unknown | 28.55 (45.17) | 31.07 (46.29) | 24.72 (43.16) |
| **Any nurse** | | | |
| Black | 21.08 (40.79) | 24.51 (43.03) | 17.61 (38.11) |
| White | 62.33 (48.46) | 55.93 (49.66) | 68.37 (46.52) |
| Race unknown | 32.63 (46.89) | 34.53 (47.56) | 28.88 (45.34) |

Notes: N=3,236 for full sample, N=1,677 for non-Hispanic Black sample, N=1,056 for non-Hispanic White sample. Sample of index births who were assigned to the treatment group. "Full sample" refers to all index births who were assigned to the treatment group. "Black" refers to Black, non-Hispanic index birth mothers who were assigned to the treatment group. "White" refers to White, non-Hispanic index birth mothers who were assigned to the treatment group. For "Main nurse" and "Any nurse", "Black" refers to Black nurses without restriction on Hispanic ethnicity due to missingness. "White" refers to White nurses without restriction on Hispanic ethnicity due to missingness. Standard errors in parentheses.

Table 9. Racial concordance analysis of NFP visits and screening

| | Participation in program | | Avg. visit length | | Visit count | | Num. dep. screens | | Any dep. screen | | Any MH referral | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Black Mother | -0.04 | -0.04 | -0.62 | -0.99 | -1.09* | -1.03** | -0.15* | -0.16* | 0.01 | 0.02 | -0.02 | -0.02 |
| | (0.02) | (0.02) | (0.94) | (0.82) | (0.43) | (0.40) | (0.07) | (0.07) | (0.02) | (0.01) | (0.02) | (0.02) |
| | | | | | | | | | | | | |
| Black Nurse | -0.10* | | -2.54* | | -1.64* | | -0.11 | | -0.03 | | -0.04 | |
| | (0.04) | | (1.22) | | (0.69) | | (0.14) | | (0.03) | | (0.03) | |
| | | | | | | | | | | | | |
| Black Mother*Black Nurse | 0.12* | 0.10* | -1.39 | 0.55 | 2.18** | 1.61* | 0.33* | 0.31 | 0.05 | 0.04 | -0.01 | -0.02 |
| | (0.05) | (0.05) | (1.59) | (1.43) | (0.79) | (0.80) | (0.16) | (0.16) | (0.03) | (0.03) | (0.04) | (0.04) |
| Adjusted $R^2$ | 0.040 | 0.111 | 0.131 | 0.314 | 0.142 | 0.265 | 0.123 | 0.222 | 0.026 | 0.162 | 0.080 | 0.222 |
| Mean | 0.74 | 0.74 | 62.23 | 62.23 | 14.26 | 14.26 | 2.45 | 2.45 | 0.92 | 0.92 | 0.14 | 0.14 |
| Mean (White Mother, White Nurse) | 0.76 | 0.76 | 62.65 | 62.65 | 15.09 | 15.09 | 2.48 | 2.48 | 0.92 | 0.92 | 0.16 | 0.16 |
| P-value | 0.45 | | 0.00 | | 0.22 | | 0.01 | | 0.15 | | 0.02 | |
| Nurse FE | | Yes | | Yes | | Yes | | Yes | | Yes | | Yes |

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

Notes: N=1,951. Sample restricted to NFP treatment mothers of non-Hispanic ethnicity, of White or Black race, whose most frequent nurse is of Black or White race. "Black Mother" refers to Black, non-Hispanic mothers. "Black Nurse" refers to Black nurses without restriction on Hispanic ethnicity due to missingness. Variables are measured across pregnancy and up to 60 days postpartum. "Participation in program' is whether the mother received NFP visits up to 60 days postpartum. "Avg. visit length" is the average (mean) NFP visit length in minutes for completed visits. "Visit count" is the number of at home NFP completed visits. "P-value" is for the test of "Black Mother, Black Nurse" = "Black Mother, White Nurse". All models includes controls for IA, teen mother, mother aged 28 and over, daily interaction with father, less than high school education, work status, receipt of social service programs, housing insecurity, high risk pregnancy. Robust standard errors in parentheses.

# 7 Discussion and conclusion

In this analysis of the impact of a randomised control trial of the Nurse Family Partnership in South Carolina on maternal mental health we observe high rates of depression or anxiety during pregnancy through 60 days postpartum with approximately 30% of control group mothers receiving an outpatient diagnosis. This is likely an underestimate of the true prevalence as many mothers may be undiagnosed. This substantial mental health burden does not correspond to high rates of treatment; only 13% of control group mothers receive medication and 5% attend therapy. This treatment gap could be driven by a number of factors: stigma and shame around seeking treatment, limited access to treatment due to maternity and psychiatric care deserts (Wisner et al., 2024), or limited engagement by providers or lack of resources to engage with the referral and follow-up phase (Valdes et al., 2023; Byatt et al., 2012). Approaches that seek to improve perinatal mental health through increased access to screening and referral to existing resources alone may be insufficient to address and improve mental health outcomes.

We find that NFP has no statistically significant effects on mental health diagnoses or treatment in a sample of first-time mothers who were income-eligible for Medicaid. This may be because there are limits to nurses' ability to positively impact maternal mental health. NFP is broad in its scope; mental health is only one of the topics covered in visits. According to home visitors, it is also one of the most challenging and sensitive topics to address, along with substance use and intimate partner violence (Duggan et al., 2018). One of the attractive features of NFP is the personal relationship that a nurse can build with a mother, but this form of social support may be insufficient in addressing maternal depression and anxiety during the perinatal period. A more specific intervention targeting maternal mental health may be more effective.

We observe substantial disparities in mental health outcomes by race in our sample. In the control group, non-Hispanic Black mothers were less likely than non-Hispanic White mothers to be diagnosed with a mental health condition (24% vs. 43%) and receive treatment (10% vs. 21% for medication, 4% vs. 6% for therapy). Participation in NFP does not appear to alleviate these disparities. Black mothers assigned to NFP were less likely to receive medication for a mental health condition relative to control Black mothers (-3 percentage points), whereas White treatment mothers were more likely to

receive medication relative to White control mothers (+4 percentage points). These differences, however, are not robust to multiple hypothesis testing correction and should be further explored in future research.

We explore mechanisms that could explain differences in Black and White mothers' experiences with the NFP program through analysis of pre-existing mental health conditions, intensity of participation in NFP, and mental health screening and referrals through both NFP and routine medical care. Differences in pre-existing symptoms cannot explain the diagnosis and treatment disparities, nor visit duration or frequency given that the difference in visits is substantively small. We do not find large racial differences in screening rates, although nurses were more likely to adhere to screening for White mothers. Increased referrals for White mothers could be a relevant factor. NFP nurses were more likely to refer White mothers for mental health services than Black mothers (20% versus 14%). Our findings contrast those in Dodge et al. (2022); the authors find statistically significant reductions in racial disparities in maternal anxiety and depression from participation in a home visiting program. Our results may differ because of program differences. The authors study "Family Connects", a short-term home visiting program delivered only during the postpartum period that was offered universally. Overall, we need more evidence on the effects of nurse home visiting in large samples of Black mothers.

The racial disparities observed in our medical claims and survey data, both in terms of baseline levels and differences in response to the treatment, are a function of maternal service utilisation, mothers' revelation of information to healthcare providers, physician behaviour and beliefs, depression screening tools, and the accuracy of reporting. There may be racial differences in the willingness to report depressive symptoms due to fear of negative repercussions. Black mothers may be worried about being considered a poor mother, or being reported to child welfare services if symptoms are reported. Black families are more likely to be reported to child welfare services (Krase, 2013), judged unfit parents, and have their children taken away and placed in foster care relative to White mothers (Yi et al., 2020). Black and White mothers may have different preferences for mental health treatment; Leis et al. (2011) document that Black mothers fear that providers will jump to prescribing unwanted and unnecessary medication. Prior research also finds that Black

patients have concerns as to whether mental health providers can culturally relate to them, particularly in the setting of therapy (Jones et al., 2015; Leis et al., 2011). Symptoms of depression and anxiety differ by race, so providers may not notice subtle or uncommon presentations of depression and anxiety (Bailey et al., 2019). Black mothers may use different language to describe depression (Baker, 2001), and may be more likely to report somatic rather than affective symptoms (Lara-Cinisomo et al., 2020). Further, few studies have tested the performance of depression screening tools among racially diverse groups of pregnant and postpartum women in the U.S. (Ukatu et al., 2018; Tandon et al., 2012). Additionally, nurses and medical providers may unconsciously or consciously be treating Black and White mothers differently, due to implicit bias or prejudice. There is a large literature documenting racial bias in the U.S. (e.g. Dehon et al. (2017); Williams and Wyatt (2015)). Screening, diagnosis, and treatment for depression and anxiety could all be influenced by provider bias. Lastly, we may not observe mental health treatment that mothers are receiving in our data. Services provided at community centres, such as peer support groups, will not appear in medical claims data. Black mothers may be referred to these services at a higher rate than White mothers, including by NFP, for many reasons (including those discussed above).

While racial concordance between patient and provider has been found to reduce racial disparities in other settings (Greenwood et al., 2020), we do not see strong evidence that racial concordance benefits Black mothers in our setting. In fact, Black mothers who see a Black nurse are significantly less likely to receive a mental health referral than White mothers who see a White nurse. A limitation of our concordance analysis is that a small percentage of our sample represents Black mothers seen by Black nurses, limiting power. We also lack sufficient data to be able to track the race of all nurses. Despite the robustness of estimates to the inclusion of nurse fixed effects, we caution against interpreting these results as causal, as unobservable confounders may influence mother-nurse pairings and affect our estimates.

Future work could seek to isolate the role of specific mechanisms discussed in this paper to better understand the mental health diagnosis and treatment gap between Black and White mothers. For example, whether there are racial differences in preferences for depression treatment, and whether healthcare providers exhibit racial biases towards mothers with depression. Research

that explores Black mothers' perceptions about the potential risks of revealing information to providers of intensive home visiting programs would also be a valuable contribution.

# References

**Alsan, Marcella, Owen Garrick, and Grant Graziani**, "Does Diversity Matter for Health? Experimental Evidence from Oakland," *American Economic Review*, 2019, *109* (12), 4071–4111.

**Baicker, Katherine, Mary Ann Bates, Annetta Zhou, Michelle Woodford Martin, and Margaret McConnell**, "Randomized Evaluation of the Nurse Family Partnership in South Carolina," 2021. AEA RCT Registry, January 20.

**Bailey, Rahn Kennedy, Josephine Mokonogho, and Alok Kumar**, "Racial and Ethnic Differences in Depression: Current Perspectives," *Neuropsychiatric Disease and Treatment*, 2019, pp. 603–609.

**Baker, F.M.**, "Diagnosing Depression in African Americans," *Community Mental Health Journal*, 2001, *37*, 31–38.

**Beeber, Linda, Alasia Ledford, Mariarosa Gasbarro, Paula Zeanah, Michael Knudtson, Sharon Sprinkle, Georgette McMichael, Alison Mosqueda, and Jennifer Leeman**, "Developing a Multicomponent Implementation Strategy for Mental Health Interventions Within the Nurse-Family Partnership: An Application of the EPIS Framework," *Journal of Nursing Scholarship*, 2022, *54* (4), 445–452.

**Benjamini, Yoav and Yosef Hochberg**, "Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society: Series B (Methodological)*, 1995, *57* (1), 289–300.

**Biaggi, Alessandra, Susan Conroy, Susan Pawlby, and Carmine M Pariante**, "Identifying the Women at Risk of Antenatal Anxiety and Depression: A Systematic Review," *Journal of Affective Disorders*, 2016, *191*, 62–77.

**Boyd-Swan, Casey, Chris M Herbst, John Ifcher, and Homa Zarghamee**, "The earned income tax credit, mental health, and happiness," *Journal of Economic Behavior & Organization*, 2016, *126*, 18–38.

**Bütikofer, Aline, Julie Riise, and Meghan M. Skira**, "The impact of paid maternity leave on maternal health," *American Economic Journal: Economic Policy*, 2021, *13* (1), 67–105.

**Byatt, Nancy, Tiffany A Moore Simas, Rebecca S Lundquist, Julia V Johnson, and Douglas M Ziedonis**, "Strategies for Improving Perinatal Depression Treatment in North American Outpatient Obstetric settings," *Journal of Psychosomatic Obstetrics & Gynecology*, 2012, *33* (4), 143–161.

**Campbell, Frances, Gabriella Conti, James J Heckman, Seong Hyeok Moon, Rodrigo Pinto, Elizabeth Pungello, and Yi Pan**, "Early Childhood Investments Substantially Boost Adult Health," *Science*, 2014, *343* (6178), 1478–1485.

**Case, Anne, Darren Lubotsky, and Christina Paxson**, "Economic Status and Health in Childhood: The Origins of the Gradient," *American Economic Review*, 2002, *92* (5), 1308–1334.

**Casey, Patrick, Susan Goolsby, Carol Berkowitz, Deborah Frank, John Cook, Diana Cutts, Maureen M Black, Nieves Zaldivar, Suzette Levenson, Tim Heeren, and Alan Meyers**, "Maternal Depression, Changing Public Assistance, Food Security, and Child Health Status," *Pediatrics*, 2004, *113* (2), 298–304.

**Catherine, Nicole, Michael Boyle, Yufei Zheng, Lawrence McCandless, Hui Xie, Rosemary Lever, Debbie Sheehan, Andrea Gonzalez, Susan M Jack, Amiram Gafni, Lil Tonmyr, Lenora Marcellus, Colleen Varcoe, Ange Cullen, Kathleen Hjertaas, Caitlin Riebe, Nikolina Rikert, Ashvini Sunthoram, Ronald Barr, Harriet MacMillan, and Charlotte Waddell**, "Nurse Home Visiting and Prenatal Substance Use in a Socioeconomically Disadvantaged Population in British Columbia: Analysis of Prenatal Secondary Outcomes in an Ongoing Randomized Controlled Trial," *Canadian Medical Association Open Access Journal*, 2020, *8* (4), E667–E675.

**Conti, Gabriella, James J Heckman, and Rodrigo Pinto**, "The Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviour," *The Economic Journal*, 2016, *126* (596), F28–F65.

_ , **Joyce Smith, Elizabeth Anson, Susan Groth, Michael Knudtson, Andrea Salvati, and David Olds**, "Early Home Visits and Health Outcomes in Low-Income Mothers and Offspring: 18-Year Follow-Up of a Randomized Clinical Trial," *JAMA Network Open*, 2024, *7* (1), e2351752–e2351752.

**Currie, Janet and Douglas Almond**, "Human capital development before age five," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1315–1486.

**Dagher, Rada K, Hannah E Bruckheim, Lisa J Colpe, Emmeline Edwards, and Della B White**, "Perinatal Depression: Challenges and Opportunities," *Journal of Women's Health*, 2021, *30* (2), 154–159.

**Deave, Toity, Jon Heron, Jonathan Evans, and Albert Emond**, "The Impact of Maternal Depression in Pregnancy on Early Child Development,"

*BJOG: An International Journal of Obstetrics & Gynaecology*, 2008, *115* (8), 1043–1051.

**Dehon, Erin, Nicole Weiss, Jonathan Jones, Whitney Faulconer, Elizabeth Hinton, and Sarah Sterling**, "A Systematic Review of the Impact of Physician Implicit Racial Bias on Clinical Decision Making," *Academic Emergency Medicine*, 2017, *24* (8), 895–904.

**Dodge, Kenneth A, W Benjamin Goodman, Yu Bai, Debra L Best, Peter Rehder, and Sherika Hill**, "Impact of a Universal Perinatal Home-Visiting Program on Reduction in Race Disparities in Maternal and Child Health: Two Randomised Controlled Trials and a Field Quasi-Experiment," *The Lancet Regional Health–Americas*, 2022, *15*.

**Duggan, Anne, Ximena A Portilla, Jill H Filene, Sarah Shea Crowne, Carolyn J Hill, Helen Lee, and Virginia Knox**, "Implementation of Evidence-Based Early Childhood Home Visiting: Results from the Mother and Infant Home Visiting Program Evaluation. OPRE Report 2018-76A.," *Office of Planning, Research and Evaluation*, 2018.

**Eckenrode, John, Mary I Campa, Pamela A Morris, Charles R Henderson Jr, Kerry E Bolger, Harriet Kitzman, and David L Olds**, "The Prevention of Child Maltreatment Through the Nurse Family Partnership Program: Mediating Effects in a Long-Term Follow-Up Study," *Child Maltreatment*, 2017, *22* (2), 92–99.

**Ettner, Susan L**, "New evidence on the relationship between income and health," *Journal of health economics*, 1996, *15* (1), 67–85.

**Evans, William N and Craig L Garthwaite**, "Giving mom a break: The impact of higher EITC payments on maternal health," *American economic Journal: economic policy*, 2014, *6* (2), 258–290.

**Frakes, Michael D and Jonathan Gruber**, "Racial Concordance and the Quality of Medical Care: Evidence from the Military," Technical Report, National Bureau of Economic Research 2022.

**Frank, Richard G and Ellen Meara**, "The Effect of Maternal Depression and Substance Abuse on Child Human Capital Development," Technical Report, National Bureau of Economic Research 2009.

**Gangopadhyaya, Anuj, Fredric Blavin, Breno Braga, and Jason Gates**, "Credit where it is due: Investigating pathways from earned income tax credit expansion to maternal mental health," *Health Economics*, 2020, *29* (9), 975–991.

**Getahun, Darios, Yinka Oyelese, Morgan Peltier, Meiyu Yeh, Vicki Y Chiu, Harpreet Takhar, Nehaa Khadka, Nana Mensah, Chantal Avila, and Michael J Fassett**, "Trends in Postpartum Depression by Race/Ethnicity and Pre-Pregnancy Body Mass Index," *American Journal of Obstetrics & Gynecology*, 2023, *228* (1), S122–S123.

**Gourevitch, Rebecca, Chloe Zera, Michelle W Martin, R. Annetta Zhou, Mary Ann Bates, Katherine Baicker, and Margaret McConnell**, "Home Visits With a Registered Nurse Did Not Affect Prenatal Care In A Low-Income Pregnant Population: Study Examines Registered Nurse Home Visits and Prenatal Care among Medicaid Enrollees," *Health Affairs*, 2023, *42* (8), 1152–1161.

**Goyal, Deepika, Caryl Gay, and Kathryn Lee**, "How Much Does Low Socioeconomic Status Increase the Risk of Prenatal and Postpartum Depressive Symptoms in First-Time Mothers?," *Women's Health Issues*, 2010, *20* (2), 96–104.

**Greenwood, Brad, Rachel Hardeman, Laura Huang, and Aaron Sojourner**, "Physician-Patient Racial Concordance and Disparities in Birthing Mortality for Newborns," *Proceedings of the National Academy of Sciences*, 2020, *117* (35), 21194–21200.

**_ , Seth Carnahan, and Laura Huang**, "Patient-Physician Gender Concordance and Increased Mortality among Female Heart Attack Patients," *Proceedings of the National Academy of Sciences*, 2018, *115* (34), 8569–8574.

**Gresenz, Carole Roan, Roland Sturm, Lingqi Tang et al.**, "Income and mental health: Unraveling community and individual level relationships," *Journal of Mental Health Policy and Economics*, 2001, *4* (4), 197–204.

**Grossman, Daniel S, Sebastian Tello-Trillo, and Barton Willage**, "Health Insurance for Whom? The 'Spill-up'Effects of Children's Health Insurance on Mothers," Technical Report, National Bureau of Economic Research 2022.

**Guldi, Melanie and Sarah Hamersma**, "The effects of pregnancy-related Medicaid expansions on maternal, infant, and child health," *Journal of Health Economics*, 2023, *87*, 102695.

**Heckman, James, Rodrigo Pinto, and Peter Savelyev**, "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes," *American economic review*, 2013, *103* (6), 2052–2086.

**Hill, Andrew J, Daniel B Jones, and Lindsey Woodworth**, "Physician-Patient Race-Match Reduces Patient Mortality," *Journal of Health Economics*, 2023, *92*, 102821.

**Huybrechts, Krista F, Kristin Palmsten, Helen Mogun, Mary Kowal, Jerry Avorn, Soko Setoguchi-Iwata, and Sonia Hernández-Díaz**, "National Trends in Antidepressant Medication Treatment Among Publicly Insured Pregnant Women," *General Hospital Psychiatry*, 2013, *35* (3), 265–271.

**Jacobs, Francine, Ann Easterbrooks, Jayanthi Mistry, Erin Bumgarner, Rebecca Fauth, Jessica Goldberg, Jessica Greenstone, Maryna Raskin, and Mariah Contreras**, "The Massachusetts Healthy Families Evaluation-2 (MHFE-2): A Randomized, Controlled Trial of a Statewide Home Visiting Program for Young Parents," *Medford, Massachusetts*, 2015.

**Johnston, David W, Stefanie Schurer, and Michael A Shields**, "Exploring the intergenerational persistence of mental health: Evidence from three generations," *Journal of health economics*, 2013, *32* (6), 1077–1089.

**Jones, Lani V, Laura Hopson, Lynn Warner, Eric R Hardiman, and Tana James**, "A Qualitative Study of Black Women's Experiences in Drug Abuse and Mental Health Services," *Affilia*, 2015, *30* (1), 68–82.

**Kaiser Family Foundation**, "Medicaid Coverage for Women," https://www.kff.org/womens-health-policy/issue-brief/medicaid-coverage-for-women/ 2022. Accessed:2025-06-27.

**Kennedy-Moulton, Kate, Sarah Miller, Petra Persson, Maya Rossin-Slater, Laura Wherry, and Gloria Aldana**, "Maternal and Infant Health Inequality: New Evidence from Linked Administrative Data," Technical Report, National Bureau of Economic Research 2022.

**Kim, Young**, "Reps. Young Kim, Robin Kelly Lead Resolution Recognizing Maternal Mental Health Awareness week," https://rb.gy/af7tge 2024. Online; accessed 16 May 2024.

**Kingston, Dawn, Maureen Heaman, Deshayne Fell, and Beverley Chalmers**, "Comparison of Adolescent, Young Adult, and Adult Women's Maternity Experiences and Practices," *Pediatrics*, 2012, *129* (5), e1228–e1237.

**Kozhimannil, Katy and Helen Kim**, "Maternal Mental Illness," *Science*, 2014, *345* (6198), 755–755.

___ , **Connie Mah Trinacty, Alisa B Busch, Haiden A Huskamp, and Alyce S Adams**, "Racial and Ethnic Disparities in Postpartum Depression Care among Low-Income Women," *Psychiatric Services*, 2011, *62* (6), 619–625.

**Krase, Kathryn S**, "Differences in Racially Disproportionate Reporting of Child Maltreatment Across Report Sources," *Journal of Public Child Welfare*, 2013, *7* (4), 351–369.

**Lara-Cinisomo, Sandraluz, Tanitoluwa Demilade Akinbode, and Jayme Wood**, "A Systematic Review of Somatic Symptoms in Women with Depression or Depressive Symptoms: Do Race or Ethnicity Matter?," *Journal of Women's Health*, 2020, *29* (10), 1273–1282.

**Leis, Julie A, Tamar Mendelson, Deborah F Perry, and S Darius Tandon**, "Perceptions of Mental Health Services Among Low-Income, Perinatal African-American Women," *Women's Health Issues*, 2011, *21* (4), 314–319.

**Lenhart, Otto**, "The effects of state-level earned income tax credits on suicides," *Health economics*, 2019, *28* (12), 1476–1482.

**McConnell, Margaret A, R Annetta Zhou, Michelle W Martin, Rebecca A Gourevitch, Maria Steenland, Mary Ann Bates, Chloe Zera, Michele Hacker, Alyna Chien, and Katherine Baicker**, "Protocol for a Randomized Controlled Trial Evaluating the Impact of the Nurse-Family Partnership's Home Visiting Program in South Carolina on Maternal and Child Health Outcomes," *Trials*, 2020, *21*, 1–21.

___ , ___ , ___ , ___ , ___ , ___ , ___ , ___ , ___ , **and** ___ , "Pre-Analysis Plan: Impact of Nurse-Family Partnership on Maternal and Early Child Outcomes," https://www.hsph.harvard.edu/sc-nfp-study/wp-content/uploads/sites/2613/2021/02/Pre-Analysis-Plan_Maternal-and-Early-Child-Outcomes.pdf 2021. Online; accessed 13 May 2024.

___ , **Slawa Rokicki, Samuel Ayers, Farah Allouch, Nicolas Perreault, Rebecca A Gourevitch, Michelle W Martin, R Annetta Zhou, Chloe Zera, Michele R Hacker et al.**, "Effect of an intensive nurse home visiting program on adverse birth outcomes in a medicaid-eligible population: a randomized clinical trial," *JAMA*, 2022, *328* (1), 27–37.

**Mejdoubi, Jamila, Silvia CCM van den Heijkant, Frank JM van Leerdam, Martijn W Heymans, Remy A Hirasing, and Alfons AM Crijnen**, "Effect of Nurse Home Visits vs. Usual Care on Reducing Intimate Partner Violence in Young High-Risk Pregnant Women: a Randomized Controlled Trial," *PloS one*, 2013, *8* (10), e78185.

_ , _ , _ , **Matty Crone, Alfons Crijnen, and Remy A HiraSing**, "Effects of Nurse Home Visitation on Cigarette Smoking, Pregnancy Outcomes and Breastfeeding: a Randomized Controlled Trial," *Midwifery*, 2014, *30* (6), 688–695.

**Michalopoulos, Charles, Kristen Faucetta, Carolyn J Hill, XA Portilla, Lori Burrell, Helen Lee, Anne Duggan, Virginia Knox et al.**, "Impacts on Family Outcomes of Evidence-Based Early Childhood Home Visiting: Results From the Mother and Infant Home Visiting Program Evaluation," *Research and Evaluation Administration for Children and Families US Department of Health and Human Services*, 2019, *7*.

**Molina, Abigail Palmer, Dorian E Traube, and Allison Kemner**, "Addressing Maternal Mental Health to Increase Participation in Home Visiting," *Children and Youth Services Review*, 2020, *113*, 105025.

**Olds, David L**, "The nurse–family partnership: An evidence-based preventive intervention," *Infant Mental Health Journal*, 2006, *27* (1), 5–25.

_ , **Harriet Kitzman, Elizabeth Anson, Joyce A Smith, Michael D Knudtson, Ted Miller, Robert Cole, Christian Hopfer, and Gabriella Conti**, "Prenatal and Infancy Nurse Home Visiting Effects on Mothers: 18-year Follow-Up of a Randomized Trial," *Pediatrics*, 2019, *144* (6).

_ , _ , **Robert Cole, JoAnn Robinson, Kimberly Sidora, Dennis W Luckey, Charles R Henderson Jr, Carole Hanks, Jessica Bondy, and John Holmberg**, "Effects of Nurse Home-Visiting on Maternal Life Course and Child Development: Age 6 Follow-Up Results of a Randomized Trial," *Pediatrics*, 2004, *114* (6), 1550–1559.

_ , **JoAnn Robinson, Lisa Pettitt, Dennis W Luckey, John Holmberg, Rosanna K Ng, Kathy Isacks, Karen Sheff, and Charles R Henderson Jr**, "Effects of Home Visits by Paraprofessionals and By Nurses: Age 4 Follow-Up Results of a Randomized Trial," *Pediatrics*, 2004, *114* (6), 1560–1568.

**O'Brien, Ruth A, Patricia Moritz, Dennis W Luckey, Maureen W McClatchey, Erin M Ingoldsby, and David L Olds**, "Mixed Methods Analysis of Participant Attrition in the Nurse-Family Partnership," *Prevention Science*, 2012, *13*, 219–228.

**O'Connor, Elizabeth, Caitlyn A Senger, Michelle Henninger, Bradley N Gaynes, Erin Coppola, and Meghan Soulsby Weyrich**, "Interventions to Prevent Perinatal Depression: A Systematic Evidence Review for the US Preventive Services Task Force," 2019.

**Perry, Cynthia D**, "Does Treating Maternal Depression Improve Child Health Management? The Case of Pediatric Asthma," *Journal of Health Economics*, 2008, *27* (1), 157–173.

**Persson, Petra and Maya Rossin-Slater**, "When dad can stay home: fathers' workplace flexibility and maternal health," *American Economic Journal: Applied Economics*, 2024, *16* (4), 186–219.

**Propper, Carol, John Rigg, and Simon Burgess**, "Child health: evidence on the roles of family income and maternal mental health from a UK birth cohort," *Health economics*, 2007, *16* (11), 1245–1269.

**Puma, Michael, Robert Olsen, Stephen Bell, and Cristofer Price**, "What to Do when Data Are Missing in Group Randomized Controlled Trials," https://ies.ed.gov/ncee/pdf/20090049.pdf 2009. Online; accessed 13 May 2024.

**Rokicki, Slawa, Dea Oviedo, Nicolas Perreault, Chloe Zera, Alecia McGregor, Mary Ann Bates, R. Annettta Zhou, Katherine Baicker, and Margaret McConnell**, "Home Visits and the Use of Routine and Emergency Postpartum Care among Low Income Women: A Randomized Clinical Trial," *JAMA Network Open*, 2024.

**Sandner, Malte, Thomas Cornelissen, Tanja Jungmann, and Peggy Herrmann**, "Evaluating the Effects of a Targeted Home Visiting Program on Maternal and Child Health Outcomes," *Journal of Health Economics*, 2018, *58*, 269–283.

**Schmidt, Lucie, Lara Shore-Sheppard, and Tara Watson**, "The effect of safety net generosity on maternal mental health and risky health behaviors," *Journal of Policy Analysis and Management*, 2023, *42* (3), 706–736.

**Sidebottom, Abbey, Marc Vacquier, Elizabeth LaRusso, Darin Erickson, and Rachel Hardeman**, "Perinatal Depression Screening Practices in a Large Health System: Identifying Current State and Assessing Opportunities to Provide More Equitable Care," *Archives of Women's Mental Health*, 2021, *24*, 133–144.

**Song, Dayoung, Roberta G Sands, and Yin-Ling Irene Wong**, "Utilization of Mental Health Services by Low-Income Pregnant and Postpartum Women on Medical Assistance," *Women & Health*, 2004, *39* (1), 1–24.

**South Carolina Department of Health and Human Services**, "Screening, Brief Intervention, and Referral to Treatment," https://www.scdhhs.gov/sites/default/files/SBIRT%20Provider%20Presentation.pdf 2010. Online; accessed 23 July 2024.

___, "Physicians Services Provider Manual," https://provider.scdhhs.gov/internet/pdf/manuals/Physicians/Manual.pdf 2024. Online; accessed 2 December 2024.

**Steenland, Maria, Dea Oviedo, Mary Ann Bates, R. Annetta Zhou, Chloe Zera, Katherine Baicker, and Margaret McConnell**, "Effect of an Intensive Nurse Home Visiting Program on Postpartum Contraceptive Use and Birth Spacing: A Randomized Controlled Trial," *ObstetGynecol*, 2024.

**Tandon, S Darius, Fallon Cluxton-Keller, Julie Leis, Huynh-Nhu Le, and Deborah F Perry**, "A Comparison of Three Screening Tools to Identify Perinatal Depression Among Low-Income African American women," *Journal of Affective Disorders*, 2012, *136* (1-2), 155–162.

**The American College of Obstetricians and Gynecologists**, "Implementing Perinatal Mental Health Screening," https://www.acog.org/programs/perinatal-mental-health/implementing-perinatal-mental-health-screening 2024. Online; accessed 31 May 2024.

**Ukatu, Nneamaka, Camille A Clare, and Mary Brulja**, "Postpartum Depression Screening Tools: a Review," *Psychosomatics*, 2018, *59* (3), 211–219.

**United States Preventive Services Task Force**, "Perinatal Depression: Preventive Interventions," https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/perinatal-depression-preventive-interventions 2019. Online; Accessed 1 May 2024.

**Valdes, Elise G, Lora Sparkman, Rola Aamar, Leigh Steiner, Jack M Gorman, Vicki Ittel, Johnny J Bethea, and Christopher Reist**, "Improving Maternal Mental Health: Assessing the Extent of Screening and Training About Peripartum Depression," *The Journal of Maternal-Fetal & Neonatal Medicine*, 2023, *36* (1), 2155042.

**Williams, David R and Ronald Wyatt**, "Racial Bias in Health Care and Health: Challenges and Opportunities," *JAMA*, 2015, *314* (6), 555–556.

**Wisner, Katherine L, Caitlin Murphy, and Megan M Thomas**, "Prioritizing Maternal Mental Health in Addressing Morbidity and Mortality," *JAMA Psychiatry*, 2024, *81* (5), 521–526.

**Yi, Youngmin, Frank R Edwards, and Christopher Wildeman**, "Cumulative Prevalence of Confirmed Maltreatment and Foster Care Placement

for U.S. Children by Race/Ethnicity, 2011-2016," *American Journal of Public Health*, 2020, *110* (5), 704–709.

**Zimmerman, Frederick J and Wayne Katon**, "Socioeconomic status, depression disparities, and financial strain: what lies behind the income-depression relationship?," *Health economics*, 2005, *14* (12), 1197–1215.

# Supplemental Appendix: The Impact of Intensive Nurse Home Visiting on Pregnancy and Postpartum Mental Health for Low-Income First-Time Mothers [*]

Abigail Dow[1][†]       Dea Oviedo [2]

Michelle Woodford Martin[3]       Mary Ann Bates[4]

Katherine Baicker[5]       Alecia McGregor[2]

Margaret McConnell[2]

July 21, 2025

[*][1] Boston University, [2] Harvard University, [3] Blue Cross Blue Shield, [4] California Cradle-To-Career Data System, [5] The University of Chicago
[†]abidow@bu.edu

# A Appendix Tables

Appendix Table 1. Sample definition

| Sample measure | Data sources | Definition |
|---|---|---|
| Index birth | South Carolina vital records birth certificates and baseline survey | A birth identified by a matched birth certificate in Vital Records with a date of birth within 120 days before or after the estimated delivery date on the baseline survey. |

Appendix Table 2. Diagnosis codes for depression, anxiety, and stress-reactions

| Description | ICD-10 code(s) |
|---|---|
| Depression | F53; O906; F99; O9934*; F32*; F33* |
| Anxiety | F41* |
| Stress-reaction | F43* |

Source: Sherman & Ali (2018)

Appendix Table 3. Diagnosis codes for substance use

| Description | ICD-10 code(s) |
|---|---|
| Neonatal abstinence disorder | P96.1, P04.2, P04.3, P04.4*, P04.8*, P04.9, P04.14, P04.16 |
| Opioids | F11* |
| Tobacco | F17.2*; O99.33*; Z72.0 |
| Alcohol | F10.* |
| Sedative | F13* |
| Cocaine | F14* |
| Amphetamines | F15* |
| Cannabis | F12* |

Source: Jarlenski et al. (2020)

Appendix Table 4. Diagnosis codes for maternal experience of violence or homicide

| Description | ICD-10 code(s) |
|---|---|
| Adult neglect | T74.01, T76.01 |
| Adult physical abuse | T74.11, T76.11 |
| Adult sexual abuse | T74.21, T76.21 |
| Adult psychological abuse | T74.31, T76.31 |
| Unspecified adult maltreatment | T74.91, T76.91 |
| Husband, perpetrator of maltreatment | Y0701 |
| Assault by unarmed brawl | Y040 |
| Observation after rape | Z0441 |
| Assault by other bodily force | Y048 |
| Unspecified perpetrator of maltreatment | Y079 |
| Other family member perpetrator of maltreatment | Y07499 |
| Encounter for mental health services for victim of spousal or partner abuse | Z691* |
| Encounter for mental health services for victim or perpetrator of other abuse | Z698* |
| Encounter for observation following alleged adult physical abuse | Z0471 |

Sources: Davidov et al. (2017), Schafer et al. (2008)

Appendix Table 5. Therapeutic class codes for filled antidepressants or anxiolytics

| Description | Therapeutic class codes |
|---|---|
| Antidepressants | 281604 |
| Anxiolytics | 2824* |

Appendix Table 6. Outpatient psychotherapy codes

| Description | CPT codes |
|---|---|
| Psychotherapy (individual, family, or group) | 90804-90815; 90832-90834; 90836- 90840; 90845-90847; 90849; 90853; 90857; 90862; 90875; 90876 |

Source: Sherman & Ali (2018)

Appendix Table 7. Balance table, by race

| | Black | | | White | | |
|---|---|---|---|---|---|---|
| | Treatment | Control | Diff. | Treatment | Control | Diff. |
| Variables | Mean | Mean | P-val | Mean | Mean | P-val |
| Age (%) | | | | | | |
| - 15-18 | 16.95 | 14.66 | 0.14 | 18.50 | 20.08 | 0.45 |
| - 19-24 | 57.36 | 58.98 | 0.43 | 52.91 | 52.39 | 0.84 |
| - 25-34 | 23.40 | 25.18 | 0.32 | 25.90 | 25.81 | 0.97 |
| - 35+ | 2.29 | 1.18 | 0.06 | 2.68 | 1.72 | 0.24 |
| BMI (%) | | | | | | |
| - <18.5 | 4.28 | 4.54 | 0.77 | 7.99 | 6.81 | 0.41 |
| - 18.5-24.9 | 35.99 | 34.36 | 0.43 | 38.91 | 44.36 | 0.04 |
| - 25-29.9 | 22.00 | 23.07 | 0.55 | 19.83 | 18.09 | 0.41 |
| - >30.0 | 37.73 | 38.04 | 0.88 | 33.27 | 30.74 | 0.31 |
| Education (%) | | | | | | |
| - < HS | 18.64 | 16.88 | 0.28 | 25.67 | 25.10 | 0.80 |
| - HS | 38.80 | 37.46 | 0.51 | 33.21 | 31.61 | 0.52 |
| - < BA | 34.98 | 36.15 | 0.56 | 34.42 | 38.12 | 0.15 |
| - BA+ | 7.52 | 9.51 | 0.09 | 6.70 | 5.17 | 0.24 |
| Mental health/substance use (%) | | | | | | |
| - Depressive symptoms | 21.07 | 21.33 | 0.88 | 16.82 | 17.12 | 0.88 |
| - High stress | 66.55 | 65.83 | 0.72 | 66.35 | 65.44 | 0.72 |
| - MH treatment | 8.04 | 7.70 | 0.77 | 24.28 | 23.61 | 0.77 |
| - Alcohol use | 50.83 | 53.16 | 0.27 | 54.41 | 51.63 | 0.30 |
| - Cigarette use | 16.80 | 18.61 | 0.26 | 42.12 | 42.77 | 0.80 |
| Other health indicators (%) | | | | | | |
| - ED use | 55.66 | 55.63 | 0.99 | 48.84 | 51.43 | 0.33 |
| - Fair/poor health | 12.25 | 8.90 | 0.01 | 14.03 | 14.04 | 1.00 |
| Family planning (%) | | | | | | |
| - FP/birth control access | 59.58 | 59.38 | 0.93 | 58.63 | 59.08 | 0.86 |
| - Desire for children | 62.87 | 62.25 | 0.76 | 71.14 | 74.38 | 0.17 |
| Baseline pregnancy | | | | | | |
| - Gestational age | 14.98 | 14.56 | 0.11 | 14.36 | 14.42 | 0.85 |
| - Started prenatal care | 85.14 | 84.53 | 0.69 | 85.00 | 82.82 | 0.26 |
| Socioeconomic indicators (%) | | | | | | |
| - Used social services | 72.27 | 73.42 | 0.54 | 59.72 | 63.69 | 0.13 |
| - Worked for pay | 54.60 | 54.96 | 0.86 | 54.63 | 50.77 | 0.15 |
| - Lived with parents | 49.74 | 50.83 | 0.60 | 36.26 | 38.31 | 0.43 |
| - Housing insecure | 12.91 | 12.78 | 0.93 | 21.13 | 23.71 | 0.24 |
| Observations | 1,705 | 846 | 2,551 | 1,081 | 523 | 1,604 |

Notes: Survey data collected after enrolment and before randomisation between 2016 and 2020. "Black" refers to non-Hispanic Black mothers. "White" refers to non-Hispanic White mothers. BMI is body mass index. '<HS' is less than a high school diploma. "HS" is a high school diploma or equivalent. "<BA" is some college, less than a bachelor's degree. "BA+" is a bachelor's degree or higher. "Depressive symptoms" is a Patient Health Questionnaire 2 score of $>= 3$. " High stress" is a Perceived Stress Scale 4 score of $>=4$. "MH treatment" is receipt of mental health treatment in the last year. "Alcohol use" is reported alcohol use in the 3 months before pregnancy. "Cigarette use" is reported cigarette use in the 3 months before pregnancy. "ED use" is use of the emergency department in the 6 months before enrolment. "Fair/poor health" is self-reported health as fair or poor. "FP/birth control access" is whether the mother has access to family planning or birth control. "Desire for children" is self-reported desire for children in the future. "Gestational age" is at enrolment in weeks. "Started prenatal care" is receipt of at least one prenatal visit before enrolment. Used social services" is use of programmes such as TANF; SNAP; SSI; WIC; and unemployment benefits. "Housing insecure" is whether the mother moved at least twice in the previous 12 months.

Appendix Table 8. Adjusting for Multiple Hypothesis Testing

|  | Unadjusted P-value | Q-value |
|---|---|---|
| **Full sample** | | |
| Any OP diagnosis/medication/therapy | 0.28 | 0.64 |
| Any OP diagnosis | 0.20 | 0.61 |
| Any Medication | 0.95 | 0.97 |
| Any Therapy | 0.50 | 0.89 |
| Any IP/ED visit | 0.12 | 0.61 |
| Num. IP/ED visits | 0.20 | 0.61 |
| Treatment follow up | 0.86 | 0.97 |
| Any substance use | 0.66 | 0.97 |
| Any domestic violence | 0.97 | 0.97 |

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Full sample N =4,932. "Q-value" is calculated using the Benjamini and Hochberg (1995) procedure. 'Any OP diagnosis/medication/therapy" is any outpatient diagnosis for depression or anxiety, medication, or therapy. "Any IP/ED visit" is an inpatient or emergency department visit with a postpartum depression or anxiety diagnosis, excluding labour and delivery. "Num. IP/ED" is defined the same, but counts the number of visits. "Treatment follow up" is a second prescription or therapy visit within 120 days of first prescription or therapy visit.

Appendix Table 9. Adjusting for multiple hypothesis testing, by race

|  | Unadjusted P-value | Q-value |
|---|---|---|
| **Black** | | |
| Any OP diagnosis/medication/therapy | 0.30 | 0.97 |
| Any OP diagnosis | 0.72 | 0.97 |
| Any Medication | 0.02 | 0.21 |
| Any Therapy | 0.77 | 0.97 |
| Any IP/ED visit | 0.65 | 0.97 |
| Num. IP/ED visits | 0.39 | 0.97 |
| Treatment follow up | 0.63 | 0.97 |
| Any substance use | 0.95 | 0.97 |
| Any domestic violence | 0.97 | 0.97 |
| **White** | | |
| Any OP diagnosis/medication/therapy | 0.77 | 0.77 |
| Any OP diagnosis | 0.45 | 0.57 |
| Any Medication | 0.04 | 0.38 |
| Any Therapy | 0.13 | 0.38 |
| Any IP/ED visit | 0.10 | 0.38 |
| Num. IP/ED visits | 0.39 | 0.57 |
| Treatment follow up | 0.33 | 0.57 |
| Any substance use | 0.54 | 0.61 |
| Any domestic violence | 0.44 | 0.57 |

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

Notes: Non-Hispanic White sample N= 1,604. Non-Hispanic Black sample N= 2,551. "OP" stands for outpatient. "L&D" stands for labour and delivery. "IP" stands for inpatient. "ED" stands for emergency department. "Any OP diagnosis/medication/therapy" is any outpatient diagnosis for depression or anxiety, medication, or therapy. "Any IP/ED visit" is an inpatient or emergency department visit with a postpartum depression or anxiety diagnosis, excluding labour and delivery. "Num. IP/ED" is defined the same, but counts the number of visits. "Treatment follow up" is a second prescription or therapy visit within 120 days of first prescription or therapy visit.

Appendix Table 10. Treatment effects of NFP, vulnerable subsample

| | Treatment Mean | Control Mean | Unadjusted ITT | | ITT |
|---|---|---|---|---|---|
| Any OP diagnosis/medication/therapy | 39.09 | 40.23 | -1.14 | -1.41 | [-5.12, 2.30] |
| Any OP diagnosis | 35.21 | 37.58 | -2.37 | -2.73 | [-6.41, 0.96] |
| Any Medication | 17.57 | 16.65 | 0.93 | 0.92 | [-2.11, 3.95] |
| Any Therapy | 8.25 | 8.45 | -0.20 | 0.29 | [-1.90, 2.48] |
| Any IP/ED visit | 18.41 | 16.27 | 2.14 | 1.58 | [-1.42, 4.59] |
| Num. IP/ED visits | 0.34 | 0.31 | 0.03 | 0.03 | [-0.07, 0.12] |
| Treatment follow up | 13.51 | 13.62 | -0.11 | 0.37 | [-2.32, 3.05] |
| Any substance use | 24.81 | 26.61 | -1.80 | -1.70 | [-4.95, 1.54] |
| Any domestic violence | 2.39 | 1.51 | 0.88 | 1.00 [+] | [-0.14, 2.13] |

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

Notes: N=2,466. Vulnerable sample comprises of those younger than 19 years, who had not finished high school, or who had challenges with mental health (PHQ-2 score of at least 3 at baseline or reported receiving mental health treatment in the year before enrolment). For the "Any" variables, the means and effects are percentages. For the "Num" variable, the means and effects are numbers. "OP" stands for outpatient. "L&D" stands for labour and delivery. "IP" stands for inpatient. "ED" stands for emergency department. "Any OP diagnosis/medication/therapy" is any outpatient diagnosis for depression or anxiety, medication, or therapy. "Any IP/ED visit" is an inpatient or emergency department visit with a postpartum depression or anxiety diagnosis, excluding labour and delivery. "Num. IP/ED" is defined the same, but counts the number of visits. "Any treatment follow up" is a second prescription or therapy visit within 120 days of first prescription or therapy visit. "Any treatment" is any medication or therapy visit. "Any therapy follow up" is a second therapy visit within 120 days of first therapy visit. "Any medication follow up" is a second medication fill within 120 days of first medication fill. 95% confidence intervals in brackets. All models control for implementing NFP agency, age groups, race and ethnicity, gestational age at time of study enrolment, relationship with father of the child, education, employment, receipt of social services, housing stability, health status, psychological state, access and utilisation of health care, alcohol and smoking use, stress, pregnancy risk factors, family planning metrics, and dummies for missing baseline covariates.

Appendix Table 11. Treatment effects of NFP, subsample of mothers with any Medicaid coverage

| | Treatment Mean | Control Mean | Unadjusted ITT | | ITT |
|---|---|---|---|---|---|
| Any OP diagnosis/medication/therapy | 32.22 | 33.71 | -1.49 | -1.56 | [-4.05, 0.92] |
| Any OP diagnosis | 28.80 | 30.48 | -1.67 | -1.75 | [-4.17, 0.67] |
| Any Medication | 13.37 | 13.41 | -0.04 | -0.04 | [-1.96, 1.88] |
| Any Therapy | 5.07 | 4.72 | 0.35 | 0.38 | [-0.81, 1.57] |
| Any IP/ED visit | 14.02 | 12.35 | 1.67 | 1.49 | [-0.36, 3.35] |
| Num. IP/ED visits | 0.25 | 0.21 | 0.03 | 0.03 | [-0.02, 0.08] |
| Treatment follow up | 8.98 | 8.81 | 0.16 | 0.07 | [-1.49, 1.63] |
| Any substance use | 21.70 | 22.66 | -0.95 | -0.51 | [-2.63, 1.61] |
| Any domestic violence | 1.63 | 1.68 | -0.05 | -0.04 | [-0.81, 0.72] |

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: N=4,864. Sensitivity analysis conducted on sample of mothers with any Medicaid coverage up to 60 days postpartum. For the "Any" variables, the means and effects are percentages. For the "Num" variable, the means and effects are numbers. "OP" stands for outpatient. "L&D" stands for labour and delivery. "IP" stands for inpatient. "ED" stands for emergency department. "Any OP diagnosis/medication/therapy" is any outpatient diagnosis for depression or anxiety, medication, or therapy. "Any IP/ED visit" is an inpatient or emergency department visit with a postpartum depression or anxiety diagnosis, excluding labour and delivery. "Num. IP/ED" is defined the same, but counts the number of visits. "Any treatment follow up" is a second prescription or therapy visit within 120 days of first prescription or therapy visit. "Any treatment" is any medication or therapy visit. "Any therapy follow up" is a second therapy visit within 120 days of first therapy visit. "Any medication follow up" is a second medication fill within 120 days of first medication fill. 95% confidence intervals in brackets. All models control for implementing NFP agency, age groups, race and ethnicity, gestational age at time of study enrolment, relationship with father of the child, education, employment, receipt of social services, housing stability, health status, psychological state, access and utilisation of health care, alcohol and smoking use, stress, pregnancy risk factors, family planning metrics, and dummies for missing baseline covariates.

Appendix Table 12. Treatment effects of NFP, subsample of mothers with continuous Medicaid coverage

| | Treatment Mean | Control Mean | Unadjusted ITT | | ITT |
|---|---|---|---|---|---|
| Any OP diagnosis/medication/therapy | 32.41 | 33.93 | -1.51 | -1.31 | [-3.83, 1.20] |
| Any OP diagnosis | 29.00 | 30.63 | -1.63 | -1.45 | [-3.91, 1.00] |
| Any Medication | 13.40 | 13.44 | -0.04 | 0.07 | [-1.88, 2.01] |
| Any Therapy | 5.07 | 4.76 | 0.32 | 0.39 | [-0.81, 1.59] |
| Any IP/ED visit | 14.16 | 12.37 | 1.79 | 1.78 [+] | [-0.10, 3.66] |
| Num. IP/ED visits | 0.25 | 0.21 | 0.04 | 0.04 | [-0.01, 0.09] |
| Treatment follow up | 8.96 | 8.88 | 0.08 | 0.08 | [-1.49, 1.66] |
| Any substance use | 21.64 | 22.70 | -1.06 | -0.39 | [-2.53, 1.75] |
| Any domestic violence | 1.66 | 1.71 | -0.05 | -0.04 | [-0.82, 0.75] |

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

Notes: N=4,770. Sensitivity analysis conducted on sample of mothers with continuous Medicaid coverage up to 60 days postpartum. For the "Any" variables, the means and effects are percentages. For the "Num" variable, the means and effects are numbers. "OP" stands for outpatient. "L&D" stands for labour and delivery. "IP" stands for inpatient. "ED" stands for emergency department. "Any OP diagnosis/medication/therapy" is any outpatient diagnosis for depression or anxiety, medication, or therapy. "Any IP/ED visit" is an inpatient or emergency department visit with a postpartum depression or anxiety diagnosis, excluding labour and delivery. "Num. IP/ED" is defined the same, but counts the number of visits. "Any treatment follow up" is a second prescription or therapy visit within 120 days of first prescription or therapy visit. "Any treatment" is any medication or therapy visit. "Any therapy follow up" is a second therapy visit within 120 days of first therapy visit. "Any medication follow up" is a second medication fill within 120 days of first medication fill. 95% confidence intervals in brackets. All models control for implementing NFP agency, age groups, race and ethnicity, gestational age at time of study enrolment, relationship with father of the child, education, employment, receipt of social services, housing stability, health status, psychological state, access and utilisation of health care, alcohol and smoking use, stress, pregnancy risk factors, family planning metrics, and dummies for missing baseline covariates.

Appendix Table 13. Treatment-on-the-treated effects of NFP

|  | Treatment Mean | Control Mean | | TOT |
|---|---|---|---|---|
| Any OP diagnosis/medication/therapy | 31.81 | 33.17 | -1.82 | [-5.27, 1.62] |
| Any OP diagnosis | 28.44 | 29.99 | -2.12 | [-5.48, 1.25] |
| Any Medication | 13.20 | 13.19 | 0.15 | [-2.51, 2.80] |
| Any Therapy | 5.01 | 4.64 | 0.57 | [-1.09, 2.22] |
| Any IP/ED visit | 13.93 | 12.34 | 2.00 | [-0.59, 4.58] |
| Num. IP/ED visits | 0.24 | 0.21 | 0.04 | [-0.02, 0.11] |
| Treatment follow up | 8.86 | 8.67 | 0.27 | [-1.88, 2.43] |
| Any substance use | 21.49 | 22.42 | -0.62 | [-3.56, 2.32] |
| Any domestic violence | 1.67 | 1.65 | 0.10 | [-0.95, 1.15] |

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

Notes: N=4,770. "TOT" is the treatment-on-the-treated effect, instrumenting NFP visits to 60 days postpartum with randomisation into the NFP treatment group. For the "Any" variables, the means and effects are percentages. For the "Num" variable, the means and effects are numbers. "OP" stands for outpatient. "L&D" stands for labour and delivery. "IP" stands for inpatient. "ED" stands for emergency department. "Any OP diagnosis/medication/therapy" is any outpatient diagnosis for depression or anxiety, medication, or therapy. "Any IP/ED visit" is an inpatient or emergency department visit with a postpartum depression or anxiety diagnosis, excluding labour and delivery. "Num. IP/ED" is defined the same, but counts the number of visits. "Any treatment follow up" is a second prescription or therapy visit within 120 days of first prescription or therapy visit. "Any treatment" is any medication or therapy visit. "Any therapy follow up" is a second therapy visit within 120 days of first therapy visit. "Any medication follow up" is a second medication fill within 120 days of first medication fill. 95% confidence intervals in brackets. All models control for implementing NFP agency, age groups, race and ethnicity, gestational age at time of study enrolment, relationship with father of the child, education, employment, receipt of social services, housing stability, health status, psychological state, access and utilisation of health care, alcohol and smoking use, stress, pregnancy risk factors, family planning metrics, and dummies for missing baseline covariates.

Appendix Table 14. Mental health status prior to intervention, by race

|  | Depressive Symptoms | High Stress | Received MH Treatment |
|---|---|---|---|
| Black | 0.05*** | 0.01 | -0.15*** |
|  | (0.02) | (0.02) | (0.02) |
| Adjusted $R^2$ | 0.02 | 0.02 | 0.07 |
| Mean | 0.20 | 0.65 | 0.14 |

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

Notes: N=2,786. Survey data collected after enrolment and before randomisation between 2016 and 2020. "Black" refers to non-Hispanic Black mothers. "Depressive Symptoms" is a Patient Health Questionnaire 2 score of at least 3. "High Stress" is a Perceived Stress Scale 4 score of at least 4. All models include controls for IA, teen mother, mother aged 28 and over, daily interaction with father, less than high school education, work status, receipt of social service programmes, housing insecurity, high risk pregnancy. Robust standard errors in parentheses.

Appendix Table 15. Retention in NFP, by race

|  | NFP to birth | NFP to 60 days pp | NFP to 1 year pp |
|---|---|---|---|
| Black | 0.03 | -0.00 | -0.03 |
|  | (0.02) | (0.02) | (0.02) |
| Adjusted $R^2$ | 0.03 | 0.04 | 0.07 |
| Mean | 0.78 | 0.70 | 0.52 |

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

Notes: N=2,786. Survey data collected after enrolment and before randomisation between 2016 and 2020. "Black" refers to non-Hispanic Black mothers. "NFP to birth" denotes that the mother remained in the NFP programme until birth. "NFP to 60 days pp" denotes that the mother remained in the NFP programme up to 60 days postpartum. "NFP to 1 year pp" denotes that the mother remained in the NFP programme up to 1 year postpartum. All models includes controls for IA, teen mother, mother aged 28 and over, daily interaction with father, less than high school education, work status, receipt of social service programmes, housing insecurity, high risk pregnancy. Robust standard errors in parentheses.

Appendix Table 16. NFP visits, by race

|  | Visit length Prenatal | Visit length 60 days pp | Visit length Study period | Visit count Prenatal | Visit count 60 days pp | Visit count Study period |
|---|---|---|---|---|---|---|
| Black | -2.44** | -2.34 | -2.04** | -0.18 | -0.45*** | -0.41 |
|  | (0.83) | (1.69) | (0.73) | (0.21) | (0.11) | (0.30) |
| Observations | 2,733 | 2,105 | 2,734 | 2,742 | 2,144 | 2,743 |
| Adjusted $R^2$ | 0.11 | 0.04 | 0.12 | 0.08 | 0.15 | 0.11 |
| Mean | 65.89 | 62.02 | 64.83 | 9.792 | 4.914 | 13.63 |

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

Notes: Survey data collected after enrolment and before randomisation between 2016 and 2020. "Black" refers to non-Hispanic Black mothers. "60 days pp" denotes that the visit took place after delivery, up to 60 days postpartum. "Study period" includes the prenatal period, delivery, and up to 60 days postpartum. "Avg. visit length" in minutes. "Visit count" is the number of at home NFP completed visits. All models include controls for IA, teen mother, mother aged 28 and over, daily interaction with father, less than high school education, work status, receipt of social service programmes, housing insecurity, high risk pregnancy. Robust standard errors in parentheses.

Appendix Table 17. NFP mental health screening and referrals, by race

| | Screened at intake | | | Screened at 36w | | | Screened 1-8w pp | | | Num. screenings | | | Any screening | | | Any MH referral | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) |
| Black | 0.02 | 0.03* | 0.03* | -0.04* | -0.02 | -0.03 | -0.02 | 0.00 | -0.01 | -0.08 | 0.02 | -0.02 | 0.02 | 0.03* | 0.02 | -0.04** | -0.04** | -0.04* |
| | (0.01) | (0.01) | (0.01) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.05) | (0.05) | (0.05) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| Adjusted $R^2$ | 0.00 | 0.04 | 0.04 | 0.00 | 0.03 | 0.04 | 0.00 | 0.04 | 0.06 | 0.00 | 0.15 | 0.16 | 0.00 | 0.02 | 0.02 | 0.00 | 0.03 | 0.03 |
| Mean | 0.87 | 0.87 | 0.87 | 0.60 | 0.60 | 0.60 | 0.66 | 0.66 | 0.66 | 2.56 | 2.56 | 2.56 | 0.90 | 0.90 | 0.90 | 0.13 | 0.13 | 0.13 |
| Controls | No | IA | Full | No | IA | Full | No | IA | Full | No | IA | Full | No | IA | Full | No | IA | Full |

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

Notes: N=2,786. Survey data collected after enrolment and before randomisation between 2016 and 2020. "Black" refers to non-Hispanic Black mothers. "Screened at 36w" is screened at 36 weeks pregnant. "Screened 1-8w pp" is screened 1-8 weeks after birth. "Any MH referral" is ever referred for mental health treatment. "IA" controls are controls for the NFP implementating agency (IA). "Full" controls include controls for the IA, teen mother, mother aged 28 and over, daily interaction with father, less than high school education, work status, receipt of social service programmes, housing insecurity, high risk pregnancy. Robust standard errors in parentheses.


Appendix Table 18. Treatment effects of NFP on Medicaid screening and referrals, by race

| | White | | Black | | Difference |
|---|---|---|---|---|---|
| | Control Mean | ITT (1) | Control Mean | ITT (2) | (1)-(2) P |
| Any depression screening | 22.56 | -0.33 [-4.76, 4.10] | 26.48 | -2.57 [-6.08, 0.93] | 0.43 |
| Any SBIRT screening | 28.49 | 4.02 + [-0.68, 8.71] | 31.68 | -0.12 [-3.91, 3.67] | 0.17 |
| Any depression or SBIRT screening | 44.93 | 2.56 [-2.68, 7.79] | 49.05 | -0.03 [-4.14, 4.09] | 0.44 |
| Any SBIRT referral | 10.52 | 0.76 [-2.33, 3.85] | 6.50 | 0.41 [-1.55, 2.37] | 0.85 |

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

Notes: Non-Hispanic White sample N= 1604. Non-Hispanic Black sample N= 2551. "ITT" is the intent to treat effect. For all variables, the means and ITT are percentages. "Any depression screening" is whether the mother received a depression screen (Edinburgh Postnatal Depression Scale) during any well-child visit. "Any SBIRT screening" is whether the mother received the "Screening, Brief Intervention, and Referral to Treatment" screen. "Any SBIRT referral" is whether the mother was referred to a specialist from an SBIRT screen. 95% confidence intervals in brackets. All models control for implementing NFP agency, age groups, race and ethnicity, gestational age at time of study enrolment, relationship with father of the child, education, employment, receipt of social services, housing stability, health status, psychological state, access and utilisation of health care, alcohol and smoking use, stress, pregnancy risk factors, family planning metrics, and dummies for missing baseline covariates.

Appendix Table 19. Medicaid coverage, by race

|  | Black (1) | White (2) | (1)-(2) |
|---|---|---|---|
| Variables | % | % | P |
| Any Medicaid coverage | 100.00 | 100.00 | . |
| Continuous Medicaid coverage | 98.46 | 98.10 | 0.38 |

Notes: Non-Hispanic White sample N= 1,579. Non-Hispanic Black sample N= 2,531. "Black" refers to non-Hispanic Black mothers. "White" refers to non-hispanic White mothers. "Any Medicaid coverage" is whether the mother had any Medicaid coverage up to 60 days postpartum. "Continuous Medicaid coverage" is whether the mother had continuous Medicaid coverage up to 60 days postpartum.