Deprivation Payments, Regional Disparities in Birth Outcomes, and the Role of Community Midwives

EsCHER Working Paper No. 2025001
January 2025

Joaquim Vidiella-Martin
Tom Van Ourti
Loes Bertens

Erasmus University Rotterdam
Making Minds Matter

Title
Deprivation Payments, Regional Disparities in Birth Outcomes, and the Role of Community Midwives

Authors
Joaquim Vidiella-Martin, University of Oxford and Erasmus Centre for Health Economics Rotterdam
Tom Van Ourti, Erasmus School of Health Policy and Management, Erasmus School of Economics, and Erasmus Centre for Health Economics Rotterdam
Loes Bertens, Department of Obstetrics and Gynaecology, Erasmus Medical Centre, and Erasmus Centre for Health Economics Rotterdam

Corresponding author: Joaquim Vidiella-Martin, email: joaquim.vidiellamartin@phc.ox.ac.uk

Keywords
deprivation, inequalities, salary increase, community midwives, birth outcomes.

JEL classification
I10, J08, J13.

Cite as
Vidiella-Martin, J., Van Ourti, T., Bertens, L. (2025). Deprivation Payments, Regional Disparities in Birth Outcomes, and the Role of Community Midwives. EsCHER Working Paper Series No. 2025001, Erasmus University Rotterdam. Available from: https://www.eur.nl/en/research/research-groups-initiatives/escher/research/working-papers

Erasmus Centre for Health Economics Rotterdam (EsCHER) is part of Erasmus University Rotterdam.
Want to know more about EsCHER? Visit www.eur.nl/escher
Want to contact EsCHER? E-mail escher@eur.nl

Interested in more EsCHER Working Papers? Download from www.eur.nl/escher/research/workingpapers

Deprivation Payments, Regional Disparities in Birth Outcomes, and the Role of Community Midwives

Joaquim Vidiella-Martin[1,2] Tom Van Ourti[2,3,4] Loes Bertens[2,5]

[1] *Department of Primary Care Health Sciences, University of Oxford*
[2] *Erasmus Centre for Health Economics Rotterdam, Erasmus University Rotterdam*
[3] *Erasmus School of Health Policy and Management, Erasmus School of Economics, Erasmus University Rotterdam*
[4] *Tinbergen Institute*
[5] *Department of Obstetrics and Gynaecology, Erasmus Medical Centre*

Abstract

Deprivation payments provide additional per-patient fees to health professionals serving patients in socioeconomically deprived areas. These payments financially compensate for higher workloads in deprived areas, but their effect on health gaps between deprived and non-deprived areas is unknown. We evaluate the effectiveness of this policy instrument by assessing its adoption among Dutch community midwives. Leveraging a discontinuity in payments based on regional deprivation scores, we show that deprivation payments successfully reduced the gap in birth outcomes between deprived and non-deprived areas. We find large improvements in health at birth among boys with no detectable effects for girls. These improvements reflect improved fetal growth during pregnancy but do not result from longer gestation periods. The improvements are present among low-risk boys, who are exclusively cared for by community midwives, and high-risk boys, who community midwives refer to more medicalized tiers of care, which is consistent with improvements arising from both better prenatal midwifery care and more timely referrals among high-risk cases before delivery. We further show that *(i)* mothers did not relocate to deprived areas to obtain care from practices receiving higher remuneration and *(ii)* geographical movement of midwives switching practices can, at most, account for a very small part of the treatment effects. Our finding that unconditional fee increases improve care aligns with the fair wage-effort theory, which postulates that effort increases as wage matches the perceived fair wage by employees.

Keywords: deprivation, inequalities, salary increase, community midwives, birth outcomes.
JEL Codes: I10, J08, J13.

This study was funded by the Erasmus Initiative Smarter Choices for Better Health. We are grateful to Sebastian Bauhoff, Jasper V. Been, Meltem Daysal, Pilar García Gómez, Wendy Janssen, Eeshani Kandpal, Hans van Kippersluis, Laura Hering, Maarten Lindeboom, Owen O'Donnell, Eddy Van Doorslaer, Miriam Wüst, and seminar participants at Erasmus University Rotterdam, Princeton University, University of Oxford, the University of Manchester, Universitat Pompeu Fabra, and meetings by the Spanish Health Economics Association, the International Health Economics Association, the Association for Public Policy and Management, and the American Society for Health Economics for their valuable feedback. We have no conflicts of interest to disclose. All errors are our own. Corresponding author: Joaquim Vidiella-Martin. Email: joaquim.vidiellamartin@phc.ox.ac.uk.

1 Introduction

Individuals in socioeconomically disadvantaged areas commonly experience disparities in health-care access and health, educational attainment, and labor market outcomes compared to their counterparts in more affluent regions of the same country. To alleviate these inequalities, policymakers can modify the financial incentives of medical and educational professionals to either motivate increased effort or attract them to these underserved regions.[1] Recently, there has been a notable proliferation of models like pay-for-performance, where the compensation of health and education providers is linked to their performance outcomes.[2] However, these models typically entail costly monitoring (De Walque et al., 2022).

Another less-explored approach is deprivation payments, which provide additional income to health and education professionals in deprived areas to offset higher workloads and encourage them to work in these areas. They are straightforward to implement and have low monitoring costs but do not directly link pay to performance. Widely applied in health and education, deprivation payments were introduced by the English NHS in the 1990s to compensate GPs in deprived areas for increased workloads (Hobbs, 1993; Crayford et al., 1995), with similar systems adopted in the Netherlands five years later (Verheij et al., 2001). In education, they have been used to attract high-quality teachers, such as in Indonesia (De Ree et al., 2018) and Uruguay (Cabrera and Webbink, 2020).[3]

While existing research has examined financial compensation for GPs working in socioeconomically deprived areas (Hobbs, 1993; Worrall et al., 1997; Verheij et al., 2001), it remains unclear whether these payments help reduce health disparities. In this paper, we study the effect of Dutch deprivation payments to community midwives (*verloskundige*) in diminishing geographical disparities in birth outcomes at the margin between deprived and non-deprived neighborhoods. Addressing these disparities, which have decreased in recent years but remain significant, is an issue long recognized as a priority by Dutch health policymakers.[4]

[1]See Cadena and Smith (2022); Devlin and Sarma (2008); Brosig-Koch et al. (2017) for a discussion on how providers adjust their behavior to an incentive plan.

[2]See De Walque et al. (2022); Cadena and Smith (2022); Eijkenaar et al. (2013); Eijkenaar (2012); Petersen et al. (2006) for reviews of the consequences of pay-for-performance schemes in healthcare on quality of care and Neal (2011) for the case of education.

[3]De Ree et al. (2018) find that higher pay for certified teachers in Indonesia improved teacher satisfaction and reduced outside employment but did not lead to measurable gains in student learning outcomes. Similarly, Cabrera and Webbink (2020) report that increased payments for teachers in disadvantaged schools in Uruguay successfully attracted more experienced teachers, with modest improvements in student performance, particularly in schools that reduced the share of inexperienced teachers.

[4]For example, perinatal mortality rates (i.e., deaths between 24 weeks of gestation and the seventh day of life) in deprived neighborhoods of large cities like Amsterdam or Rotterdam are 20% higher than the country average, while mortality in the wealthiest neighborhoods of such cities is 10 percent lower than the country average (Vos et al., 2014). For other references on inequalities in birth outcomes in the Netherlands, see Sheldon

The Dutch model views pregnancy and childbirth as natural processes requiring minimal medical intervention, complemented by one of the highest rates of home births in the developed world (Daysal et al., 2015). Community midwives are the primary medical professionals caring for women with uncomplicated pregnancies and serve as gatekeepers in the prenatal care system (Amelink-Verburg and Buitendijk, 2010; Evers et al., 2010; Wiegers and Hukkelhoven, 2010; van Teijlingen and McCaffery, 1987). As autonomous practitioners, community midwives handle all prenatal visits, check-ups, and deliveries for low-risk pregnancies without involvement from obstetricians/gynecologists (OB/GYNs) or clinical midwives,[5] unless a pregnancy is assessed as high-risk. During our study period (2008-2011), approximately 53% of pregnancies were low-risk, while 47% were referred to specialized caregivers for high-risk management before delivery.

In 2008, the Royal Dutch Organization of Midwives reported that community midwives needed, on average, 23 percent extra time when caring for women residing in socioeconomically deprived areas (Buisman and Gerats, 2008). In June 2008, the Dutch government announced the implementation of deprivation payments for community midwives, starting January 1, 2009 (Nederlandse Zorgautoriteit, 2009). These payments, amounting to 23 percent of the standard fee, were designed to compensate midwives for the additional time and resources required to care for women in deprived neighborhoods. Importantly, the policy was announced after most pregnancies affected by the payments had already been conceived, reducing the likelihood of anticipatory effects. Furthermore, the payments were automatically triggered by the mother's residential postcode, precluding any manipulation by midwives or mothers.[6]

We evaluate the impact of deprivation payments on disparities in birth outcomes during the first three years of the policy's implementation (2009-2011). We use a difference-in-discontinuities approach (Grembi et al., 2016) that exploits a continuous deprivation index measuring the deprivation status of a Dutch neighborhood, thereby effectively comparing deprived neighborhoods with almost-deprived neighborhoods before (2008) and after (2009-2011) the introduction of deprivation payments.[7] We focus on four key health outcomes at birth: low birth weight,

(2008); Zeitlin et al. (2013); Vos et al. (2016); Bertens et al. (2020); Vidiella-Martin et al. (2021).

[5] Clinical midwives work in hospitals under the supervision of an OB/GYN.

[6] Section 2.3 provides more detail on the announcement and potential anticipatory effects. Section 5.3 presents the empirical evidence confirming that births around the cutoff date do not influence the results, validating the assumption that the policy was unanticipated and that timing and maternal neighborhood of residence were not manipulated.

[7] The deprivation index is recalculated every four years. We limit our analysis to the first three years of the policy to avoid potential bias introduced by changes in the deprivation score calculation after 2011. Sensitivity analyses confirm the validity of the deprivation index as the treatment assignment variable, including tests for manipulation of the index and the lack of effects of treatment status on parental characteristics, which might correlate with birth outcomes.

small for gestational age (birth weight adjusted for gestational age and sex),[8] preterm birth (before 37 weeks of gestation), and low Apgar score. All analyses are conducted separately by gender due to the well-documented fragile male hypothesis, which suggests that male fetuses are more sensitive to adverse prenatal conditions (Kraemer, 2000).[9]

We find that the deprivation payments led to a sizable and statistically significant reduction in the gap in birth outcomes of male infants born in deprived versus non-deprived areas. Specifically, they helped reduce the gaps in low birth weight, small for gestational age, and suboptimal Apgar scores. We do not find significant effects for preterm births. Since the male newborns grew larger without a corresponding increase in the length of pregnancies, the observed improvements in birth weight are likely a result of improved fetal growth in utero rather than prolonged gestation. Notably, our quantile regressions show that gains in birth weight were relatively constant across the distribution, indicating that the payments improved fetal growth throughout the weight and gestational age distributions rather than solely benefiting those at the bottom. In contrast, we find no significant effects on birth outcomes for female infants. This treatment heterogeneity is consistent with the literature showing that boys are more severely affected by adverse prenatal conditions than girls (Sanders and Stoecker, 2015; Nilsson, 2017).[10]

These findings open the door to exploring how community midwives' responsibilities for risk assessment and referral may have influenced the observed improvements in male birth outcomes. Community midwives determine whether pregnancies are low- or high-risk based on obstetric factors and refer high-risk cases to secondary-tier care led by OB/GYNs and clinical midwives. We observe whether a pregnancy is low- or high-risk at two points during a pregnancy: right before delivery starts and at the end of delivery. We document no significant changes in the

[8]Small for gestational age, defined as birth weight below the 10th percentile for gestational age and sex, is considered a more precise measure of fetal growth compared to low birth weight as it accounts for gestational age and sex differences (De Onis and Habicht, 1996; World Health Organization, 1995). Lower gestational age is naturally associated with lower birth weight, but not all infants born with a birth weight below 2,500 grams experience compromised outcomes. Small for gestational age is widely used as a proxy for intrauterine growth restriction, a condition in which fetal growth is impaired due to factors such as placental insufficiency or maternal health complications (Morris et al., 2024). It is critical for identifying at-risk populations of adverse birth outcomes, including perinatal morbidity and mortality, and evaluating health inequalities and policy interventions (Schlaudecker et al., 2017; Lee et al., 2013). While preterm birth captures acute shocks during pregnancy, small for gestational age reflects longer-term cumulative impacts on fetal growth.

[9]Male fetuses grow faster than females from the earliest stages of gestation, leaving less time to compensate for adverse circumstances, which contributes to their higher vulnerability to adverse in-utero conditions (Broere-Brown et al., 2016; Bekedam et al., 2002). This heightened sensitivity is reflected in a greater risk of complications, including premature birth, stillbirth, and perinatal brain damage, compared to female fetuses (Singer et al., 1968; Naeye et al., 1971; Mizuno, 2000).

[10]Sanders and Stoecker (2015) examine fetal death rates and find that male fetuses have higher mortality in response to prenatal stressors. Nilsson (2017) investigate the impact of prenatal exposure to alcohol and note that male infants exhibit more significant negative health outcomes than females.

overall share of pregnancies classified as low versus high-risk right before the delivery starts in deprived areas. Moreover, the percentage of low-risk pregnancies that remained under community midwives' responsibility throughout delivery (around half of the deliveries categorized as low-risk when the delivery starts eventually transferred during delivery to OB/GYNs and clinical midwives) did not change after the introduction of deprivation payments, indicating that the observed improvements in birth outcomes were not driven by community midwives delaying or avoiding referrals until delivery to retain financial incentives.[11] Sensitivity analyses also confirm that the composition of parents, as proxied by observable characteristics, did not change during the study period.[12] However, we do not observe how timely a referral to high-risk care before delivery was made.

Given the distinct care pathways for low- and high-risk pregnancies, one might expect differential effects for each group: community midwives directly oversee all aspects of care for low-risk pregnancies. In contrast, for high-risk pregnancies, their role primarily involves risk assessment and timely referral. Our analysis reveals similar improvements in male birth outcomes for both low- and high-risk pregnancies. The absolute treatment effects are larger among high-risk pregnancies, with a higher baseline prevalence of adverse outcomes, but the relative treatment effects are similar across both groups. This finding underscores that the observed gains likely reflect enhanced care coordination rather than selective changes in one risk group alone.

To understand the policy's broader implications, we examine whether the observed improvements stem from changes in midwife behavior. We focus on four delivery-related outcomes that reflect the quality of midwifery care and complication management: referrals to OB/GYN and clinical midwife care during labor for initially low-risk pregnancies, emergency C-sections, the presence of meconium in the amniotic fluid, and birth trauma/perinatal asphyxia.[13] Our main finding is a reduction in the presence of meconium in low-risk male pregnancies, suggesting that community midwives managed these pregnancies more effectively by reducing stress or improving care during the later stages of pregnancy. Meconium is a direct indicator of fetal

[11]As detailed in Section 2, the fee received by community midwives is proportional to the part of the pregnancy they care for. Being the main caregiver during delivery leads to an increased fee for community midwives, compared to referring to an OB/GYN or clinical midwife just before delivery. We find no evidence of changes in the referral to OB/GYNs and clinical midwives just before or during delivery, which suggests no strategic behavior by community midwives to obtain a bigger financial incentive.

[12]We cannot examine how accurately low- and high-risk pregnancies were classified, as we only observe birth outcomes after children are born, and these outcomes are influenced by the care provided in each tier of care. However, we can use parental characteristics (determined before any contact with a community midwife) to shed light on whether the composition of each group changed due to the policy of interest.

[13]Birth trauma refers to physical injuries sustained during delivery, such as fractures or nerve damage, while perinatal asphyxia describes oxygen deprivation around the time of birth. These outcomes often share underlying causes, such as obstructed labor or delayed intervention, and are frequently analyzed together in research on severe delivery complications (Daysal et al., 2019).

distress and is associated with chronic stressors, such as placental insufficiency or acute events during labor.[14] We find no evidence of effects on referrals *during* delivery to OB/GYN care for low-risk pregnancies, which often reflect complications requiring urgent higher-tier intervention. Similarly, emergency C-sections, which indicate unanticipated complications requiring immediate surgical delivery, and birth trauma/perinatal asphyxia, capturing severe complications during labor, show no significant changes. These results suggest improvements may have occurred in unobservable areas, such as better prenatal care earlier in pregnancy or more timely referrals for high-risk cases *before* rather than *during* labor. Additionally, the lack of changes in referrals during delivery indicates that the reduction in meconium may have been achieved through improved prenatal care rather than increased reliance on higher-tier interventions.

We further show, using national data, that the relocation of community midwifery practices to deprived areas to take advantage of the higher per-patient payments was minimal during the study period and unlikely to explain our results.[15] Additionally, we find no evidence of changes in parental characteristics due to the introduction of deprivation payments, ruling out shifts in fertility patterns or strategic relocation by pregnant women as a major driver of the results.

Taken together, our findings suggest that male birth outcomes in deprived areas caught up with those in non-deprived areas during the first few years after the installment of deprivation payments. These improvements were most likely attributable to changes in community midwifery care, even though we cannot observe relevant indicators of midwifery behavior, such as the volume or length of prenatal visits or the timeliness of referrals before delivery. Further research could build on our findings in the short-run to explore whether future shifts in community midwifery care and general equilibrium effects could influence the longer-term dynamics.[16]

We make two contributions to the economic literature.[17] First, our paper adds to the broader literature studying optimal employee compensation schemes. While unconditional pay increases do not directly incentivize effort, 'efficiency wage' models suggest they can enhance effort through two mechanisms: reciprocity between employer and employee, which encourages additional effort in response to a wage premium (an 'effort premium'), and the higher cost of

[14] Timely obstetric interventions are associated with reduced risks of complications such as meconium aspiration syndrome (Yoder et al., 2002).

[15] Around 6.5% of community midwives change their practice every year. This shared remained constant before and after the introduction of the deprivation payments.

[16] We observe a non-significant pattern of decreasing point estimates over time when we estimate treatment effects separately for each year (2009, 2010, and 2011).

[17] Community midwife-led care in the Netherlands has been extensively studied in public health, showing comparable perinatal outcomes to obstetrician-led care for low-risk pregnancies (Amelink-Verburg and Buitendijk, 2010; Wiegerinck et al., 2020) and supporting the safety of planned home births under midwife supervision (Bolten et al., 2016). This paper focuses on the causal impacts of financial incentives within this model.

job loss, which discourages shirking behavior (Akerlof and Yellen, 1988; Akerlof, 1982; Akerlof and Yellen, 1990; Shapiro and Stiglitz, 1984). Most empirical evidence on unconditional wage increases comes from lab experiments (Esteves-Sorenson, 2018). Notable exceptions are De Ree et al. (2018) and Cabrera and Webbink (2020), discussed earlier, who study a salary increases for teachers in disadvantaged areas of Indonesia and Uruguay, respectively, finding small to no impacts on student outcomes. Unlike teachers, who are typically employed by schools with fixed academic schedules, community midwives in the Netherlands serve clients on a rolling basis, making relocation less feasible. Complaints about the additional efforts required to care for mothers in deprived neighborhoods led to the introduction of deprivation payments. Our findings of improved birth outcomes in this context support the fair wage-effort hypothesis, which suggests that workers adjust effort proportionally to the perceived fairness of their wage.

Second, we contribute to the literature on the potential of less medicalized interventions to affect prenatal health. Prenatal care is widely recognized as a key determinant of birth outcomes, with substantial evidence from public health and economics underscoring its importance (Rosenzweig and Schultz, 1983; Currie and Gruber, 1996; Almond et al., 2005; Currie, 2011). While medicalized interventions dominate much of this research, there is growing recognition of the role community health workers and non-medicalized care models can play in improving maternal and child health outcomes (Perry et al., 2014; Björkman Nyqvist et al., 2017). Historical work demonstrates that trained midwives significantly reduced maternal and infant mortality in Sweden and the United States during the late 1800s and early 1990s (Lazuka, 2018). Contemporary studies find that early engagement with midwives, particularly in outpatient prenatal visits, improves neonatal health by influencing maternal behaviors and knowledge (Cygan-Rehm and Karbownik, 2022; Altındağ et al., 2024). Our study demonstrates how financial incentives for community-based providers can reduce geographical inequalities in birth outcomes, bridging evidence on the importance of prenatal care with the role of community health workers in improving health.

The paper is organized as follows: Section 2 summarizes obstetric care in the Netherlands and the details of deprivation payments. Section 3 describes the data. The empirical strategy is outlined in Section 4, and results on infant health are presented in Section 5. Mechanisms behind the findings are discussed in Section 6, and Section 7 concludes.

2 Background

2.1 The Dutch Obstetric Care System

The Dutch obstetric care system is based on the conception that pregnancy, delivery, and puerperium (the postnatal 6-week period starting immediately after delivery) are all natural processes and excessively medicalized treatments should therefore be avoided (Vries et al., 2012; Daysal et al., 2015). Obstetric care is organized into three tiers according to the categorization of pregnancies in low- and high-risk (Amelink-Verburg and Buitendijk, 2010; Evers et al., 2010). The tier of care determines the type of professionals involved. The primary tier of care focuses on low-risk pregnancies. Community midwives are the lead (and often only) professionals providing care to women with uncomplicated pregnancies throughout the pregnancy and delivery (Wiegers and Hukkelhoven, 2010).[18] Pregnancies categorized as high-risk are cared for by the secondary and tertiary tiers of care. OB/GYNs and clinical midwives affiliated with general hospitals form the secondary tier of care. The tertiary care tier consists of OB/GYNs affiliated with academic hospitals (Lescure et al., 2017). The deprivation payments we study in this paper apply to community midwives working in the first tier but not to clinical midwives working in other two tiers, which we pool together.

The choice of place of delivery is closely tied to the risk categorization. Low-risk pregnant women can choose to deliver at home, a birth center (*geboortecentrum*, in Dutch), or in the hospital. Birth centers are community midwifery-managed locations that offer low-medicalized care to low-risk women during labor and birth (Hermus et al., 2017). During home and birth center deliveries, pregnant women are supervised by their community midwife. Women who have an increased obstetrical risk will be referred to a hospital and deliver their child under the supervision of OB/GYNs. Low-risk pregnant women can also choose to deliver at a hospital but will be supervised by their community midwife and not by OB/GYNs and will have to cover the extra costs of hospital delivery with out-of-pocket payments or additional insurance.[19] [20] Other pregnancy and maternity-related expenses for low- and high-risk pregnancies are fully covered by the basic compulsory insurance.[21] Most obstetric and maternity care costs do not

[18]During the study period, GPs are part of the primary tier of care in less than 1% of pregnancies, mostly in less populated areas.

[19]In 2010, the cost of hospital delivery for low-risk pregnancies was around €350.

[20]Because low-risk pregnancies can also be delivered in a hospital, we do not focus on the place of delivery in this paper. Instead, as detailed in the next section, we study the leading practitioner during delivery.

[21]All residents in the Netherlands are obliged to take basic health insurance which offers comprehensive coverage and is funded via general taxation and premiums of approximately €100 per month in 2010 (Enthoven and van de Ven, 2007; Van de Ven and Schut, 2008; Schäfer et al., 2010). The insured face a compulsory deductible, which in 2010 was set at €155 per year (Schäfer et al., 2010).

fall under deductible costs (van der Geest and Varkevisser, 2019).

In summary, the Dutch obstetric care system involves little to no cost-sharing and is designed around risk selection, encouraging low-medicalized and community-midwifery-led care for low-risk deliveries.

2.2 The Role of Community Midwives

Women are advised to contact a community midwife when they find out about their pregnancy. They can contact and choose their community midwife without a referral from the GP. Contact details are available online or via their GP, who provides a list of the community midwifery practices within the area. Community midwifery practices have a specified working area to safeguard timely care and have a community midwife on call 24/7.[22] They offer prenatal consults to low-risk pregnant women, supervise their deliveries, and provide postnatal visits to all mothers.

Community midwives' responsibilities include *(i)* providing information about the importance of risky health behaviors such as smoking or drinking during pregnancy;[23] *(ii)* assessing the risk levels of pregnancies by evaluating medical history and monitoring for the presence or occurrence of obstetrical or medical risk factors[24] — this process involves categorizing pregnancies as low- or high-risk, and, if classified as high-risk, transferring care to a higher tier (clinical midwives or OB/GYNs) at any point during pregnancy or delivery (Amelink-Verburg and Buitendijk, 2010; Evers et al., 2010) —; *(iii)* supervising the delivery of low-risk pregnancies at home, a birth center, or the hospital (Bais and Pel, 2006); and *(iv)* postnatal checkups during the first ten days after birth.

Community midwives hold 10 to 12 consultations with low-risk pregnant women, with each visit ranging from 10 to 45 minutes. The first visit usually takes place around the 8^{th} week of pregnancy (Manniën et al., 2012). After this, a community midwife will typically meet a low-risk pregnant woman every four weeks in the first half of their pregnancy and more frequently

[22] If a community midwife cannot visit a client in labor because she is assisting another client, she will call a colleague from her own or a neighboring practice to attend to her client.

[23] Before being pregnant, women can, in theory, seek information about pregnancy planning from a community midwifery practice or the general practitioner (Maas et al., 2022). Since the first contact with the community midwife does not generally happen before the 8^{th} week of pregnancy (Manniën et al., 2012), general practitioners almost always provide this information.

[24] Risk factors leading to the categorization of a high-risk pregnancy may include non-gynecological pre-existing conditions such as hypertension, alcoholism, or psychiatric disorders; and gynecological pre-existing conditions such as pelvic floor reconstructions; complications in a previous delivery such as preterm births or multiple miscarriages; or conditions diagnosed during pregnancy such as infections, gestational hypertension, or blood loss.

near the end. Standard care for low-risk women includes two ultrasound scans: one in the first term of the pregnancy to determine the due date and one anomaly scan at twenty weeks. A post-natal visit is often scheduled every other day during the first ten days after birth. It takes between fifteen minutes and one hour to complete.[25] During the post-natal period, community midwives collaborate with maternity assistants (*kraamverzorgster* in Dutch) who assist after birth for a recommended amount of 49 hours on average (minimum of 24 and maximum of 80 hours per newborn; Laureij et al. (2021)).[26]

If a community midwife works full-time, she will take care of prenatal care, delivery, and postnatal care of approximately 105 women annually on average. In 2009, 1,523 community midwives in the Netherlands were practicing in the first tier of care (Hingstman and Kenens, 2010). Of these, 1,355 (89%) (co-)owned their practice, and 168 (11%) worked for a community midwifery practice as an employee (i.e., they were employed by an independently established community midwife and did not co-own the practice). 92% of community midwives work in group practices (i.e., co-owned by several community midwives), while 8% worked alone in solo practices (Hingstman and Kenens, 2010). These shares were stable during our study period (2008-2011) (Hingstman and Kenens, 2009, 2010, 2011, 2012).

Community midwifery practices (including solo practices) receive a fixed amount per pregnancy cared for, which in 2009 was €1,333.50 per pregnancy, including delivery and postnatal care (Nederlandse Zorgautoriteit, 2008).[27] Despite this lack of price competition between practices, fees vary between pregnancies if the community midwifery practice does not care for the whole pregnancy (e.g., due to miscarriage or referral to the secondary or tertiary tier). In 2009, community midwifery practices received 37% of the full fee for care during the prenatal period, 40% for care during the delivery, and 22% for postnatal care.[28]

2.3 Neighborhoods and Deprivation Payments

Deprivation payments were paid to community midwifery practices delivering children born to women residing in deprived neighborhoods after January 1, 2009. In contrast, the same com-

[25]Mothers that deliver without complications in the hospital are generally discharged home within a few hours after birth and receive postpartum care at home (Hendrix et al., 2009).

[26]The maternity assistant is a skilled nurse with a lower secondary education degree who performs medical checks, supports breastfeeding, gives information, takes care of light household chores, prepares meals, and cares for other children if necessary (Laureij et al., 2021).

[27]All euro values in this paper are reported in 2009 prices.

[28]The remuneration for care during delivery that started under the community midwife's supervision is always the same amount, independent of delivery duration or whether the woman stays under the community midwife's care or is referred to OB/GYNs during birth. Likewise, the compensation for postnatal care is always the same, independent of the number of visits.

munity midwifery practices only received the standard fee for women living in non-deprived neighborhoods. The additional remuneration per pregnancy in a deprived area was unconditional — i.e., not subject to any requirement for additional health services delivery, such as screening tests or visits, nor was it conditional on meeting specific outcome benchmarks.

The payout amounted to 123% of the standard fee of €1,333.50 and was based on the average extra time (23%) required to care for women residing in socioeconomically deprived areas as reported in a 2008 survey of the Royal Dutch Organization of Midwives (Buisman and Gerats, 2008).[29] Deprivation payments were paid to the practices. They increased the direct financial reward for midwives who (co-)owned a community midwifery practice (89% of all community midwives), while the remaining 11% of community midwives worked as an employee (Hingstman and Kenens, 2010).

Importantly, community midwives could not anticipate or manipulate behavior in response to the policy. The announcement of the deprivation payments was made on June 30, 2008. By that time, most pregnancies leading to births eligible for the payments were already underway, making it implausible for community midwives to influence conception timing. Additionally, because the payments were directed to community midwifery practices and not mothers, there was no financial incentive for mothers to adjust their behavior in anticipation of the policy.[30]

A neighborhood (*buurt* in Dutch) is considered deprived when its deprivation score (*achterstandsindex* in Dutch) surpasses a certain cut-off value. The deprivation scores for 2008-2011 are constructed based on four measures: *(i)* the average number of addresses per square kilometer in 2007; *(ii)* the share of non-Western migrants[31] in 2007; *(iii)* average income of the residents in 2007; and *(iv)* the share of public benefits recipients among the population aged 15 to 64 in 2007.[32] Based on these scores, the cut-off point for neighborhood deprivation was determined such that around 5% of the population was considered deprived. For 2008-2011, this score was 5.09. Similar calculations were done in 2003 and 2012 to obtain the scores for 2003-2007 and 2012-2015. The deprivation status of a substantial number of neighborhoods

[29]When the community midwifery practice did not care for the whole pregnancy, payout equaled 123% of the standard sub-fee for the relevant period (prenatal, delivery, or postnatal care; see section 2.2).

[30]Consistent with this, our empirical results are robust to the inclusion or exclusion of births around the January 1, 2009, cut-off date, as we detail in Section 5.3. Moreover, the payout was automatic depending on the mother's residential postcodes in municipality registers and thus could not be manipulated by the community midwifery practice or the mother.

[31]Statistics Netherlands defines a person with a western migration background as someone from a country in Europe (excluding Turkey), North America and Oceania, or from Indonesia or Japan (Statistics Netherlands, 2024).

[32]The deprivation score equals the sum of the standardized values of the natural logarithm of these four variables (Wiegers and Devillé, 2008).

differed between periods.[33] Since these changes might reflect endogenously determined shifts in the neighborhood deprivation scores, the analysis is restricted to 2008-2011.

The Netherlands also employs a system of neighborhood-based deprivation payments for GPs. GP deprivation payments have been in place since 1996, and the deprivation score and cut-off value are identical to those used for the community midwife deprivation payments. We account for this competing policy using a difference-in-discontinuity approach, which we detail in Section 4.

3 Data

3.1 Data Sources

We combine several administrative data sources from Statistics Netherlands via a remote access facility. We use the Dutch perinatal register from 2008 to 2011, which contains detailed information on birth outcomes and complications, risk categorization, location of birth and medical professionals supervising the delivery, and maternal characteristics such as age or parity of all pregnancies over 22 weeks of gestation in the country. Using maternal identifiers, we merge these birth records with municipality registers, tax records, and nationality registers to identify respectively the neighborhood of residence at the date of conception and a vector of demographic and socioeconomic characteristics of pregnant women.[34] Neighborhood deprivation scores were calculated from the formula in Wiegers and Devillé (2008) using neighborhood-level data provided by Statistics Netherlands (Centraal Bureau voor de Statistiek, 2018).

3.2 Study Population and Main Sample

The study population consists of all live births between 1 January 2008 and 31 December 2011 in neighborhoods with deprivation scores $S \in [S_c - h, S_c + h]$, where S_c denotes the cut-off point triggering the deprivation payment for a given neighborhood. Stillbirths are excluded because the maternal identifiers, required to link to the neighborhood of residence of the mother, are not correctly registered for stillborn children.

[33] Wiegers and Devillé (2008) report 56 new and 44 no longer deprived neighborhoods out of a total of 215 deprived neighborhoods in 2008, and Devillé and Wiegers (2012) report 63 new and 70 no longer deprived neighborhoods out of a total of 224 deprived neighborhoods in 2012.

[34] Since pregnant women choose their community midwife at the start of the pregnancy, we assign pregnancies to the neighborhood women were registered at the date of conception. The date of conception is estimated by subtracting gestational age from the date of birth. Gestational age is detailed at week and day level, calculated by community midwives during the first consultation, and updated after the ultrasound scan in the first term (usually in weeks 10 to 12 of pregnancy). Dutch households must register their new addresses within five days of moving, thus minimizing the risk of measurement error when assigning women to neighborhoods.

Given the higher vulnerability of male versus female newborns to adverse in-utero conditions, all analyses are conducted separately for boys and girls. Males are more vulnerable to exposure to adverse in-utero circumstances (Almond and Mazumder, 2011; Sanders and Stoecker, 2015; Barreca and Page, 2015; Nilsson, 2017). From the very early stages of gestation onward, male fetuses grow faster than female fetuses (Broere-Brown et al., 2016), which results in a shorter time to compensate for adverse circumstances before growing into the next phases and often translates into a natural survival advantage for females to adverse prenatal conditions (Kraemer, 2000; Bekedam et al., 2002). The male fetus is also at greater risk of death or damage from almost all the obstetric catastrophes that can happen before birth and during delivery (including perinatal brain damage, cerebral palsy, congenital deformities of the genitalia and limbs, premature birth, and stillbirth; Singer et al. (1968); Naeye et al. (1971); Lavoie et al. (1998); Mizuno (2000); Bekedam et al. (2002)).

The preferred bandwidth of $h^* = 1.4$ used in the main analyses leaves a data set with 49,622 male and 47,106 female observations. We provide more details on the choice of bandwidth in Section 5. We exclude observations with incomplete information on birth weight, birth weight centile, gestational age, Apgar score, and risk categorization. The final data set covers 538 neighborhoods from 83 municipalities, consisting of 48,407 boys and 46,018 girls.[35]

3.3 Key Variables

Health at Birth

The main outcome variables are birth weight, birth weight centile, gestational age, and Apgar score.[36] Economists often focus on birth weight as a measure of health at birth. However, birth weight alone does not capture the distinction between longer gestational age and fetal growth. Gestational age reflects cumulative exposure to adverse in-utero conditions and is often related to acute, short-term stressors that result in premature birth. In contrast, fetal growth, measured through birth weight centile, captures the cumulative impact of longer-term in-utero circumstances and has been shown to have lasting implications for health at birth and later in life (Gluckman et al., 2005). We include two different measures to disentangle these factors. Gestational age is recorded at the daily level. To measure fetal growth, we construct birth weight centile, which ranks the newborn's birth weight from 1 to 100, relative to other infants

[35]The annual number of births by neighborhood in the study population is larger than in the Netherlands as a whole (Table A.1).

[36]The perinatal register also records early neonatal mortality, i.e., death between the first and seventh day of life. Since early neonatal mortality occurs, on average, in fewer than 2 out of 1,000 births (Table 1), we lack sufficient statistical power to credibly detect the effects of the deprivation payments on mortality and exclude this outcome from the main results. Baseline results for early neonatal mortality are reported in Appendix B.

with the same gestational age (in days) and sex in the reference Hoftiezer population (Hoftiezer et al., 2019). We also utilize the Apgar score, which assesses the newborn's physical condition through five indicators: heart rate, respiratory effort, muscle tone, response to stimulation, and skin coloration. The score is recorded at 1 and 5 minutes after birth, with the 5-minute score being used in our analysis. A score of 10 represents optimal health, and scores of 7 or above are generally considered healthy.

Improvements in continuous measures like birth weight or gestational age do not necessarily indicate improved health outcomes, especially depending on whether they occur at the top or bottom of the outcome distribution (Abrevaya and Dahl, 2008). Therefore, we construct four binary outcome variables using medical thresholds (Lee et al., 2013; Blencowe et al., 2019). These are: *(i)* low birth weight, indicating a birth weight below 2,500 grams; *(ii)* small for gestational age, defined as birth weight below the 10^{th} birth weight Hoftiezer centile, proxying fetal growth restriction (Eskes et al., 2017); *(iii)* preterm birth, defined as being born before 37 full weeks of gestation; and *(iv)* Apgar score below 7 (Apgar < 7).

Risk Profiles

Community midwives in the Netherlands play a critical role in categorizing pregnancy risk. They are responsible for the full prenatal care and delivery process for low-risk pregnancies and can improve birth outcomes by providing higher-quality care, more frequent or prolonged visits, offering health guidance, and ensuring smoother deliveries. For high-risk pregnancies, community midwives are responsible for early risk assessment and referring the mother to the secondary and third tier of care, consisting of OB/GYNs and clinical midwives. Consequently, community midwives' impact on birth outcomes depends heavily on their ability to correctly categorize risks and make timely referrals. Updating the risk categorization from low- to high-risk may occur at any point in the pregnancy up to the day of delivery. In the next section, we explain how we deal with transfers from low- to high-risk during delivery.

To proxy for risk categorization, we construct a binary variable, *Low Risk*, which takes the value 1 if the pregnancy remained under the responsibility of a community midwife until the start of the delivery and 0 if it was transferred to the secondary tier of care before the end of delivery implying that the delivery ended under the supervision of OB/GYNs and clinical midwives.[37]

[37]The perinatal register does not record the timing of the transfers, except whether it occurred before or after the start of the delivery.

Measures of Community Midwifery Care and Fetal Distress

In addition to birth outcomes and risk profiles, we analyze four variables related to community midwifery care, delivery complications, and fetal distress, which indirectly assess how community midwives responded to deprivation payments. Information on the length and frequency of prenatal visits or the content of visits is not available in the perinatal register. Therefore, we focus on variables that may reflect how well pregnancies were planned and managed, such as the occurrence of emergencies or complications.

First, we measure whether the community midwife remains the responsible caregiver until the end of the delivery. A transfer to the secondary tier of care during labor for these low-risk pregnancies can indicate complications, such as failure to progress or fetal distress, but it may also occur for non-complication-related reasons, such as pain relief, which community midwives cannot administer (Rowe et al., 2012). Deliveries that begin under the supervision of the secondary tier of care are, by definition, high-risk and always remain under the care of the secondary tier until the end of delivery. Hence, when a low-risk pregnancy is referred to high-risk care during delivery, we still classify it as low-risk, but we also observe the transfer to the secondary tier of care.

Second, we observe the occurrence of emergency C-sections, which indicate that specialized care became necessary due to complications that were not anticipated during pregnancy or early labor. Emergency C-sections are typically performed in response to acute fetal or maternal distress, such as sudden placental abruption or umbilical cord prolapse, and highlight potential shortcomings in antenatal risk assessment or unexpected complications during labor (National Health Service, 2019).

Third, we track the presence of meconium in the amniotic fluid, which is a direct indicator of fetal distress during late pregnancy or labor. Meconium can result from chronic stressors, such as placental insufficiency, maternal hypertension, or acute events during labor. Unlike transfers to secondary care, which may occur for non-complication-related reasons, meconium is more specifically associated with fetal stress and suboptimal in-utero conditions (Shaikh et al., 2010). Timely obstetric interventions can reduce the risk of complications such as meconium aspiration syndrome (Yoder et al., 2002).

Finally, we measure the presence of birth trauma or perinatal asphyxia, which captures severe complications during delivery. We follow Daysal et al. (2019) and treat these two related conditions as one binary indicator. These complications often arise from inadequate care, mismanagement of labor, or delays in necessary interventions (Collins and Popek, 2018; Daysal et al., 2019). Birth trauma/perinatal asphyxia is critical for assessing whether timely and appropriate management during labor contributed to improved outcomes for the infant.

Each of these measures captures unique dimensions of pregnancy and delivery management. Transfers to secondary care provide insights into the community midwife's ability to manage complications or anticipate the need for specialized care. Emergency C-sections capture acute, high-risk events requiring immediate intervention. Meconium reflects fetal stress that may develop gradually or acutely, and birth trauma or asphyxia assesses the effectiveness of delivery management in preventing severe complications. By examining these outcomes, we infer whether improved planning, earlier risk detection, or more timely interventions following the introduction of deprivation payments contributed to reducing complications and improving birth outcomes.

Demographic and Socioeconomic Characteristics

We also use demographic and socioeconomic characteristics: parity (1 if primiparity, 0 if multiparity), birth order, multiple births (binary indicator), maternal and paternal age at birth, whether the mother and father have Dutch nationality (1 if born in the Netherlands or double nationality) and standardized disposable household income in the year before the delivery.[38]

3.4 Descriptive Statistics

Table 1 provides descriptive statistics for male and female newborns in the study population (columns 1 and 2). Birth outcomes are reported in Panel A. Male babies weigh 3.4 kilograms on average, while female newborns' birth weight is slightly lower. Around 7% of births in the study population are low birth weight, 15% are small for gestational age,[39] 7-8% are preterm, and between 1 and 2% have Apgar < 7. Figure 1 displays the distribution of all four non-dichotomous outcomes.

Panel B presents summary statistics for the proxy of pregnancy risk, community midwifery/maternal

[38] Disposable income consists of total household income, including income transfers, but net of income taxes and contributions — such as social security contributions and health insurance premiums. The OECD modified equivalence scale (Organisation for Economic Co-operation and Development, 2008) is used to standardize for the household size in which the mother is registered. We use disposable income and household size of the year before the delivery, which are not affected by the survival of a newborn.

[39] The percentage of newborns that are small for gestational age, defined as being below the 10^{th} percentile of infants born of the same sex and gestational age, is larger than 10%, even when we use all the births in our sample. This is because the prescriptive birth weight charts were derived from live-born singleton infants, born to ostensibly healthy mothers after uncomplicated pregnancies and spontaneous onset of labor (Hoftiezer et al., 2019).

care indicators, and measures of fetal distress. There are no marked differences between the descriptives for male and female newborns. Slightly over half of the study population is classified as low-risk. Around half of these low-risk pregnancies are supervised by the community midwife throughout the entire delivery (*Birth ends supervised by a community midwife*).[40] 9-11% of all newborns were delivered via an emergency C-section, and during 7-8% of the deliveries, meconium was found present in the amniotic fluid. Birth trauma/perinatal asphyxia were relatively uncommon (less than 1% of births).

Panel C includes additional pregnancy and parental characteristics.[41] 48% of all mothers were pregnant for the first time (primiparous). Mothers were, on average, 30 years old at delivery, and fathers were 4 years older. 59% of the mothers had Dutch nationality. The average total standardized yearly disposable income was €20,300 - €20,400.

Comparing the characteristics of our study population (columns 1 and 2) with those in the Netherlands as a whole (columns 3 and 4) confirms that mothers living in deprived or close-to-deprived neighborhoods are at higher risk of adverse birth outcomes (Panel A). We also observe a higher presence of meconium in the amniotic fluid in the study population (Panel B). Parents in our study population are less wealthy, younger, and less likely to have Dutch nationality (Panel C).

4 Empirical Strategy

The deprivation score and threshold used to determine the neighborhoods in which community midwives receive deprivation payments (since 2009) are also used to identify the neighborhoods in which GPs receive deprivation payments (since 1996). Hence, a traditional regression discontinuity design would identify the combined effect of the policy of interest (i.e., deprivation payments for community midwifery practices) and the confounding policy (i.e., deprivation payments for GPs).[42] We implement a difference-in-discontinuities estimator that identifies the effect of the policy of interest, net from that of the confounding policy (Grembi et al., 2016).

[40] Note that only births that start supervised by a community midwife (that is, low-risk births) can end under the supervision of a community midwife.

[41] Paternal information is missing for some observations because the father ID is missing (unknown father's identity or other registration issues). There are also missing values for primiparous women, Dutch mothers, and standardized income. As detailed in the next section, our baseline estimates exclude the control variables in Panel C to maximize the available sample. We test the robustness of this methodological choice in Section 5.3, showing this has little impact on our baseline results.

[42] More formally, because of the confounding policy, the standard regression discontinuity assumption of continuity of potential outcomes along the deprivation score is not satisfied in a cross-sectional setting (Hahn et al., 2001).

Let us define Y_{it} as the birth outcome of infant i born in year t. The treatment year is 2009; that is, in $t \geq 2009$, only infants born in year t to mothers living at the time of conception in neighborhoods with a deprivation score S_i above a cut-off S_c are treated. The deprivation score, the running variable in our analysis, is time-invariant in our study period. As a result, treatment assignment is given by:

$$D_{it} = \begin{cases} 0 & \text{if } S_i < S_c \text{ and } t < 2009 \\ 0 & \text{if } S_i \geq S_c \text{ and } t < 2009 \\ 0 & \text{if } S_i < S_c \text{ and } t \geq 2009 \\ 1 & \text{if } S_i \geq S_c \text{ and } t \geq 2009 \end{cases}$$

Our baseline model fits linear regression functions to the observations distributed within a distance h on either side of S_c, both at $2008 \leq t < 2009$ and $2009 \leq t \leq 2011$. Formally, we restrict the sample to neighborhoods in the interval $S_i \in [S_c - h, S_c + h]$ and estimate:

$$\begin{aligned} Y_{it} = {} & \delta_0 + \delta_1 S_i^* + Treatment_i\,(\gamma_0 + \gamma_1 S_i^*) \\ & + Post_t(\alpha_0 + \alpha_1 S_i^*) + Post_t \cdot Treatment_i\,(\beta_0 + \beta_1 S_i^*) \\ & + \epsilon_{it} \end{aligned} \tag{1}$$

where scores are centered at the cut-off point ($S_i^* = S_i - S_c$). $Treatment_i$ is an indicator taking value one if the mother of i lives in a deprived neighborhood (that is, if $S_i^* \geq 0$). $Post_t$ takes value one in the post-treatment period (2009-2011) and zero in the pre-treatment year (2008). β_0 is our parameter of interest, capturing the change in the health gap at the cut-off point after the introduction of deprivation payments to community midwives. To maximize the available sample, we do not include additional controls or fixed effects in Equation 1. We show in Section 5.3 that the resulting estimates are not sensitive to this choice; our baseline results are very similar in magnitude and significance to those including a rich vector of individual controls. The standard errors are clustered at the neighborhood level, which is our unit of treatment. The optimal bandwidth h^* is calculated using the standard regression discontinuity techniques described in Calonico et al. (2020).[43] Under the identifying assumption that

[43] We are not aware of methods to define the optimal bandwidth selection in difference-in-discontinuity settings but show robustness to alternative bandwidth choices in Section 5.3.

the deprivation payments for GPs have a time-invariant effect on the outcome, the coefficient β_0 is the difference-in-discontinuities estimator, which identifies the treatment effect of community midwives' deprivation payments. We provide evidence in Section 5.3 supporting this assumption.

5 Results

5.1 Motivating Evidence

Our empirical strategy exploits a discontinuity in pay across a deprivation score threshold before and after reform. Before estimating the treatment effects, we visually inspect the discontinuity across the deprivation score of our four main outcomes (low birth weight, small for gestational age, preterm birth, and low Apgar score) in Figures A.1 and A.2, for boys and girls, respectively.[44] The graphs highlight which data points have deprivation scores in the interval $S_i \in [S_c - h, S_c + h]$, leading to the final sample of 48,407 boys and 46,018 girls described in Table A.2. The optimal bandwidths for the various outcomes according to the criterion proposed by Calonico et al. (2020) are presented in Table A.3. Our final bandwidth of choice $h^* = 1.4$ is the average of those shown in the table. Male birth outcomes, except for preterm birth, in the pre-intervention period are higher in deprived than in non-deprived neighborhoods in close vicinity to the cut-off. This pattern was reversed after 2009, suggesting that deprivation payments effectively improved health at birth among boys. We find no consistent patterns for girls.

5.2 The Effects of Deprivation Payments on Birth Outcomes

Table 2 reports the estimated effects of deprivation payments on the four primary outcomes: low birth weight, small for gestational age, preterm birth, and Apgar < 7. It also shows the effects of a single summary index to reveal the effect of deprivation payments on all outcomes jointly and to address concerns of multiple hypothesis testing.[45] We find large and statistically

[44] Corresponding graphs for the non-dichotomous outcomes are presented in Figures A.3 and A.4. For completeness, we show the same discontinuity for early neonatal mortality in Figure B.1. As discussed above, we do not find a single early neonatal death for a large part of the deprivation score bins (represented by the average mortality lying exactly at 0).

[45] Following Kling et al. (2007), we define the summary index as the sum of z-scores of its components, with the sign of each measure oriented so that more beneficial outcomes have higher scores. The z-scores are calculated by subtracting the control group mean and dividing by the standard deviation of the estimation sample. Thus, each index component has a mean of 0 and a standard deviation of 1 in our regressions. After aggregating all four z-scores, the index was standardized such that the control group has mean zero and standard deviation equal to one, thus facilitating its interpretation: an estimated treatment effect of β should be interpreted as deprivation payments causing an increase of health at birth equal to β standard deviations. In Section 5.3, we show that our findings are robust to alternative definitions of the summary index.

significant effects of deprivation payments on the birth outcomes of boys. Community midwifery deprivation payments decreased the probability of being born with low birth weight by 1.7 percentage points (relative to 6.5 percent in the sample) and small for gestational age by 3.2 percentage points (relative to an average of 14.8). Preterm birth decreased by less than 1 percentage point (but the effect size estimate is very imprecise); low Apgar score decreased by 1 percentage point (relative to an average of 2); and the composite index of health at birth increased by 0.11 standard deviations. Deprivation payments had no impact on birth outcomes of females, as none of the five outcome measures is economically or statistically significant in Panel B. The treatment effect on girls' composite index of health at birth is -0.02 (Panel B, column 5). This heterogeneity between boys and girls is aligned with economics, medical, and epidemiological evidence discussed in Section 3.2 (Almond and Mazumder, 2011; Sanders and Stoecker, 2015; Barreca and Page, 2015; Nilsson, 2017; Broere-Brown et al., 2016; Kraemer, 2000; Bekedam et al., 2002; Singer et al., 1968; Naeye et al., 1971; Lavoie et al., 1998; Mizuno, 2000).

It seems plausible that gains in birth outcomes are more likely among infants with unfavorable birth outcomes than infants who would have been healthy without deprivation payments. We explore this by estimating the effects of deprivation payments along the distribution of the *continuous* outcomes birth weight, birth weight centile, and gestational age. Estimates along the distribution of the Apgar score (and thus the summary index) are not reported as percentiles 20 to 100 all have an Apgar score of 10. As a starting step, we examine the effects of the deprivation payments on continuous birth outcomes in Table A.4 in the Appendix. Next, Figure 2 summarizes the estimates obtained from combining conditional quantile regressions with the difference-in-discontinuity design in equation 1. The results for females show no clear pattern along the distributions of the outcomes, and the null hypothesis of zero effects can not be rejected. For males, gains in birth weight are larger at higher birth weight quantiles, while correcting for gestational age shows more considerable centile gains in the central part of the birth weight centile distribution with a peak at the 30^{th} quantile, indicating that deprivation payments did have less impact at the very top and very bottom of the birth weight centile distribution. For gestational age, effects at all quantiles are not statistically distinguishable from a zero effect, although the point estimate is larger at low quantiles. Overall, we do not find evidence of the effects belonging exclusively at certain parts of the outcome distributions.

5.3 Internal Validity and Robustness Checks

This section reports several tests to evaluate the validity of the empirical approach. First, there is no evidence of manipulation in the density function of our treatment assignment variable, i.e., neighborhood-level deprivation scores. Figure A.5 indicates continuity of the density across the cut-off, and the McCrary density test ($p = 0.873$) rejects the existence of a discontinuity at the cut-off point (McCrary, 2008). Second, the identifying assumption of time-invariant discontinuities in parental characteristics around the cut-off is not rejected. Except for standardized income in the girls' sub-sample, deprivation payments' effect estimates on the full vector of parental characteristics are small and not statistically significant (see Table A.5). Third, the discontinuity at the cut-off due to community midwifery deprivation payments is identified under the assumption that the discontinuity due to GP deprivation payments is constant over time. This assumption is not testable, but we can check whether the discontinuity in birth outcomes before the introduction of community midwifery deprivation payments was time-invariant. This is only rejected in one of ten outcomes when we estimate equation 1 using data from 2006 and 2007 with the deprivation scores and cut-off that were applicable in 2004-2007 (see the coefficient estimates of the interaction terms *Treatment x 2007* in Table A.6).[46] Fourth, we test for manipulation of date of birth around the time the policy was implemented (January 1, 2009). To do so, we remove all births one week and two weeks on either side of the threshold and report the results in Tables A.7 and A.8. These show that our baseline results remain unchanged after removing observations that could be subject to manipulation of the running variable.

Four additional sensitivity tests underscore the validity of the empirical approach. A placebo test on two birth outcomes that are determined at conception (i.e., before any contact with a community midwife) — sex of the newborn and multiple pregnancies — finds economically and statistically insignificant effects (see Table A.9). Two alternative specifications for the summary index of health at birth — using Z-scores of the continuous birth outcomes in the calculation of the summary index or using the first component of a principal component analysis of the four binary outcome measures (Cygan-Rehm and Karbownik, 2022) — lead to point estimates that are very close to the baseline (see Table A.10). The inclusion of different sets of fixed effects and control variables gives virtually identical results (see Tables A.11 and A.12). Finally, varying

[46] Data for 2008 were excluded since the change in the deprivation score between 2007 and 2008 was potentially endogenous (see section 2.3), while 2004 and 2005 were not considered as the Health Insurance Act of 2006 introduced managed competition and compensated insurers for risk selection via risk-adjustment (National Health Care Institute, 2006). Community midwives' remuneration was not directly affected by the reform. Still, mothers faced different coinsurance arrangements, and the number of women entering prenatal care late, which negatively affected birth outcomes (Sonchak, 2015), decreased between 2005 and 2006.

the bandwidth from $h = h^*/2$ up to $h = 2h^*$ (with $h^* = 1.4$ in our baseline model) indicates that our estimates are relatively constant over this interval (Figure A.6).

5.4 Heterogeneity Analyses: Were All Risk Groups Affected Equally?

As detailed in Section 2, community midwives are responsible for categorizing a pregnancy risk and providing care for low-risk pregnancies. Among those low-risk pregnancies, they can improve outcomes by delivering better care (e.g., more frequent and more prolonged visits that result in better health behaviors from mothers) and by ensuring a smoother delivery. Among high-risk pregnancies, they can improve outcomes by accurately classifying the risk and referring the mother to the secondary tier of care. This illustrates how care is very much determined by risk categorization.

Because of this, we explore whether the health improvements documented above are different by risk profiles. To do so, we use our proxy for pregnancy risk, which, as described in Section 3, is whether a community midwife was the responsible care provider for a pregnancy when the delivery started. We re-estimate Equation 1 on the five baseline birth outcomes, separately by risk profiles (Table 3). Slightly more than half of the births in our sample are categorized as low-risk (Table 1). Among boys (Panels A and B), we find that the absolute effects on low birth weight and small for gestational age (columns 1 and 2) are larger for high-risk groups, which is expected since the prevalence of these two outcomes is also higher among these pregnancies. To better compare the effects across risk groups, we focus on column 5 of Table 3 (the summary index), which accounts for differences in baseline prevalence. The effects are similar for low- and high-risk boys, with estimated impacts of 0.12 (0.13) standard deviations for low-risk (high-risk) boys.[47]

Risk categorization might be directly affected by the introduction of deprivation payment. This would be the case if, for example, community midwives now categorize the marginal case (i.e., the case that lies precisely between low- and high-risk) as high-risk because they receive sufficient income from their other patients. Conversely, one could imagine the marginal case being now categorized as low-risk because the community midwife would receive more income providing the same care. Recall that, as explained in Section 2, the community midwifery fee depends on the length of the care supplied. In Table 4, we evaluate whether risk profiling changed after introducing deprivation payments. To do so, we estimate Equation 1 on our binary measure of risk profile. We find no evidence of such changes, as the estimated treatment

[47] The summary index is standardized within each risk group, allowing for direct comparison of effect sizes.

effects are small in magnitude (less than 1% relative to the sample mean for both boys and girls) and insignificant. While these groups might have had compositional changes, the total volume of cases categorized as low- and high-risk remained stable after introducing deprivation payments. In Section 6, we further delve into the care provided to each group to unravel the mechanisms behind our results.

5.5 Interpretation of Results

Our findings suggest that, following the introduction of deprivation payments in 2009, birth outcomes improved among boys in deprived areas where community midwives were receiving these payments despite the lack of explicit performance incentives. The improvements are most evident in the outcomes related to fetal growth, as measured by low birth weight and small for gestational age. This suggests that boys in the treatment group experienced better in-utero growth, holding gestational age constant, rather than longer pregnancies. Interestingly, these improvements occurred across both low- and high-risk boys despite community midwives having different responsibilities for each group. We find no improvements are found for girls.

One might be tempted to hypothesize that the community midwives' improved care drives these gains. However, other competing alternatives could explain these findings. In the next section, we provide suggestive evidence in that regard.

6 Mechanisms

This section investigates three potential mechanisms that could explain the improvements in male birth outcomes following the introduction of deprivation payments. These mechanisms are: *(i)* changes in the quality of community midwifery care, *(ii)* relocation of community midwifery practices to more deprived areas, and *(iii)* changes in the composition of mothers receiving community midwifery care, either due to shifts in fertility patterns or strategic relocation of pregnant women.

First, community midwives receiving deprivation payments may improve their care provision. They may become more diligent in their risk assessments, provide better prenatal advice, and ensure more timely referrals to secondary care for high-risk pregnancies. Second, community midwifery practices might relocate to deprived areas where the higher per-patient fees provide financial incentives, thus potentially improving access to care for mothers in these neighborhoods. Finally, changes in the composition of mothers receiving community midwifery care could result from shifts in fertility patterns or from non-deprived mothers registering in deprived areas to access community midwives receiving deprivation payments. These changes in parental characteristics may account for the observed improvements in birth outcomes, regardless of changes in community midwifery care. While this analysis sheds valuable light on how deprivation payments might have led to improved outcomes, isolating any mechanism as the primary driver behind the results remains challenging.

6.1 Changes in Community Midwifery Care

This section explores whether community midwives altered their care provision in response to the deprivation payments. Ideally, the most informative indicators of changes in community midwifery care would be the frequency and length of prenatal visits and the timing of referrals to secondary care. The frequency and length of visits could reveal how much attention community midwives provided to low-risk mothers and how they managed pregnancy-related health behaviors and conditions. Similarly, the timing of referrals (how early they assigned pregnancies to the secondary tier of care, consisting of OB/GYNs and clinical midwives) would offer insights into their caseload management and ability to detect high-risk pregnancies.

Unfortunately, this data is unavailable in our study. As a result, we focus on the next best set of observable outcomes that can provide insight into the quality of community midwifery care and how effectively they manage pregnancies. We narrow our analysis to four key outcomes: (1) a delivery ending supervised by a community midwife, which signals that no complications required transferring the birth to a more medicalized tier of care; (2) emergency C-sections (i.e., non-planned C-sections), (3) the presence of meconium in the amniotic fluid, and (4) birth trauma or perinatal asphyxia. These outcomes capture aspects of care related to risk assessment, appropriate referral timing, and managing fetal distress during delivery. While these indicators reflect events at birth, they offer indirect evidence of prenatal care quality and pregnancy management. The presence of complications during delivery (e.g., emergency C-sections, meconium, birth trauma) can be interpreted as proxies for stress during pregnancy and suboptimal fetal growth. To be consistent with our results in Section 5, we also specify a summary index (column 5).

We evaluate these results separately by risk group, as following Section 5.4. First, we examine prenatal care for low-risk pregnancies, where community midwives are responsible for prenatal care and delivery management. Panel A of Table 5 presents the results for low-risk boys. We observe a modest, non-significant increase in community midwife-led deliveries, suggesting community midwives could assess risks accurately (column 1). This finding might indicate improved risk assessment. Next, in columns 2 to 4, we examine our remaining three measures of midwifery care and fetal distress: emergency C-sections, the presence of meconium, and birth trauma/perinatal asphyxia. These outcomes capture the success of prenatal care in minimizing stress for both mother and fetus, particularly in the later stages of pregnancy. We observe a non-significant reduction in emergency C-sections and birth trauma/asphyxia, along with a significant decline in meconium, suggesting that community midwives were better able to manage pregnancy complications. These results indicate that community midwives enhanced their performance for low-risk male pregnancies. These results are backed by the summary index's positive but statistically insignificant result.

Next, we focus on high-risk boys (Panel B), where the delivery was initiated under the supervision of the secondary tier of care (OB/GYNs and clinical midwives). Since high-risk deliveries cannot be led by a community midwife, column 1 is empty for high-risk boys and girls. We observe no significant changes in the other outcomes (even if we find positive effects on health). For girls (Panels C and D), the effects across all outcomes are small, statistically insignificant, and of mixed signs.

Overall, we do not find strong evidence of widespread changes in care provision across the observable outcomes examined. The only significant finding is a reduction in the presence of meconium in low-risk male pregnancies. This may suggest that community midwives managed these pregnancies more effectively, potentially by reducing stress or improving care during the later stages of pregnancy. While this could align with a hypothesis that deprivation payments enabled community midwives to allocate more time to low-risk pregnancies — through more frequent or more prolonged visits, or by providing better guidance on managing health behaviors — this remains speculative as we do not observe direct evidence of such changes in our data.

Importantly, we do find significant improvements in birth outcomes for both low- and high-risk boys, even though we do not see corresponding changes in the cruder measures of complications like emergency C-sections or birth trauma. This suggests that the improvements may have occurred in areas we cannot directly observe, such as the frequency and quality of prenatal visits for low-risk pregnancies or the timeliness of referrals to secondary care (before delivery) for high-risk ones. It is plausible that community midwives provided more attentive care earlier in pregnancy, allowing for better management of health behaviors and earlier detection of risks. Our results are consistent with the hypothesis that midwives no longer wait to refer mothers in deprived areas to high-risk care until the delivery starts after the introduction of deprivation payments. However, this remains speculative and cannot be concluded based on our results. In the following sections, we evaluate and rule out other potential mechanisms behind our findings.

6.2 Relocation of Community Midwifery Practices

One potential driver for the observed improvements in birth outcomes is the relocation of community midwifery practices from non-deprived to deprived areas, where community midwives may seek higher remuneration due to the deprivation payments. To investigate this, we employ two strategies, both suggesting that relocation can account for only a small portion of the improvements in infant health.

First, we analyze publicly available data from the Royal Dutch Organization of Midwives' annual reports from 2009 (Hingstman and Kenens, 2010). Out of 2,444 active community midwives in the country, 80 established new practices, and the number of self-employed community midwives increased by 13. Additionally, 73 community midwives left their practice. This implies that the share of community midwives changing employment (whether through entering the labor market, moving practices, or becoming self-employed) in 2009 was approximately 6.5%. Historical data from previous years indicate that these figures were relatively stable over time (Hingstman and Kenens, 2009, 2011), suggesting that the practice changes observed in 2009 were not substantially different from other years. Assuming that only the 6.5% of community midwives who relocated are responsible for improving birth outcomes among deprived mothers (and no improvement from non-relocating community midwives), this would require the improvements in birth outcomes to be disproportionately large for relocating community midwives. Specifically, the improvements in health at birth (e.g., a reduction of 1.7 percentage points in low birth weight and 3.2 centiles in small for gestational age) would need to be about 15 times larger for relocating community midwives than for non-relocating ones, which is highly unlikely and unreasonable given the data.

Second, we examine the timing of the effects by estimating the impact of deprivation payments separately by year. Since most community midwives co-own their practices rather than being employed by another, relocation is time-consuming. As such, if relocation were a significant factor, we would expect the share of relocating community midwives to increase gradually over time, with larger effects in later years. In contrast, our results show the opposite pattern. The effects of deprivation payments were larger in the first year (2009) than in subsequent years.

In Table 6, the estimated treatment effect on the summary index for male newborns (Panel A, column 5) is 0.147 in 2009, decreasing to 0.093 in 2010 and 0.085 in 2011. Similar patterns are observed for low birth weight (Panel A, column 1) and preterm birth (Panel A, column 3), where the effects are most substantial in 2009 and taper off in later years. Moreover, we find no statistically significant effects for most outcomes related to female newborns throughout the 2009-2011 period (Panel B). Only one of the ten outcomes — small for gestational age among girls (Panel B, column 2) — rejects the null hypothesis of homogeneity of treatment effects across this period, and it only does so at the 10% significance level.

Together, these two pieces of evidence suggest that the relocation of community midwifery practices played, at most, a modest role in explaining the improvements in birth outcomes during the first three years of the reform. The consistent evidence of stronger effects in the first year (before substantial community midwife relocations could have occurred) further supports this interpretation. However, this analysis does not rule out the possibility that community midwife relocation could contribute to the longer-term impacts of the deprivation payments, which are beyond the scope of our study.

6.3 Changes in Parental Characteristics

Finally, we explore if the positive effects of deprivation payments can be explained by parents (with characteristics leading to better perinatal health) seeking the care of community midwifery practices that received deprivation payments. This would be a rational choice by parents if they expected community midwives in practices receiving deprivation payments to exert higher effort when providing obstetric care. Since community midwives can only care for pregnant mothers living near their practice, any changes in the demand for obstetric care would necessarily change residency. This, in turn, would result in different observable characteristics of mothers across both sides of the deprivation threshold before and after introducing the deprivation payments. This hypothesis can be ruled out using one of our robustness checks in Section 5.3. Table A.5 shows no significant differences in parity, birth order, parental age, nationality, or household income due to the reform. These results suggest that changes in the demand for obstetric care were not the driving factor behind our findings.

This analysis also rules out differential fertility trends over time across both sides of the deprivation threshold. If this were the driver of our results, our results in Table A.5 would show significant variation in parental characteristics.

To further confirm these results, we conduct two additional analyses. First, we use information from the municipality registers and create *Movers*, a binary indicator taking value one if the mother changed her location between the six months before conception and delivery, and zero else. Estimating equation 1 on this variable suggests this was not the case: the estimated treatment effects are very close to zero, not significant, and precisely estimated (Table 7). Second, we re-estimate our main results on a sub-sample of mothers for whom we do not observe any change in location in the six months before conception. The results of this exercise are reported in Table A.13 and are very close, both in magnitude and significance, to those of Table 2.

Together, these complementary analyses suggest that changes in parental characteristics (whether due to different fertility trends or selective relocation) are unlikely to explain the improvements in birth outcomes. This strengthens the interpretation that the observed effects are more plausibly attributed to changes in community midwifery care rather than shifts in the population composition.

7 Discussion

Reducing geographical health disparities is a priority for policymakers worldwide. Deprivation payments, which provide additional income to healthcare professionals in socioeconomically deprived areas, represent an underexplored yet promising tool for achieving this goal. Deprivation payments can improve health outcomes without imposing monitoring costs by compensating for the longer time and increased effort required to serve these populations.

This paper leverages a natural experiment in the Netherlands to investigate the short-term effects of deprivation payments to community midwives. Beginning in 2009, community midwives received a 23% fee increase for caring for mothers in deprived neighborhoods. Our results indicate that these payments helped reduce the gap in birth outcomes between deprived and non-deprived areas. Specifically, male infants in deprived neighborhoods experienced improved birth weight, birth weight centile, and Apgar scores.

While these payments effectively reduced health disparities, the policy implications depend on their underlying mechanisms. We provide suggestive evidence showing that the improvements were driven by enhanced quality of care provided by community midwives rather than mothers relocating to access better midwives or by community midwives switching practices. We observe reductions in fetal distress among low-risk male pregnancies and find no effects among high-risk male pregnancies. This may suggest that improved prenatal planning, better guidance during pregnancy, or earlier risk detection and referral to high-risk may have played a role in driving our results. These findings align with the fair wage-effort hypothesis, which posits that workers increase their effort in response to perceived wage fairness.

Our findings diverge from studies on deprivation payments in education. For instance, De Ree et al. (2018) and Cabrera and Webbink (2020) show that such payments can improve teacher satisfaction and attract experienced teachers, yet they have limited effects on student outcomes. In contrast, we show that deprivation payments to midwives improved the quality of maternal care, leading to tangible health benefits for infants.

As demand for non-medicalized maternity care rises in developed countries (Dahlen et al., 2011; Daysal et al., 2015), understanding the role of community midwives and the potential for policy-driven improvements in their care is critical. Our findings highlight the value of financial incentives in enhancing maternal care quality and open new avenues for research on optimal compensation schemes for midwives and other healthcare professionals.

Data access

The study used non-public microdata provided by Statistics Netherlands via a Remote Access facility, complies with the data agreements for the use of non-public microdata provided by Statistics Netherlands and Perined (CBS Project 8490 and Perined Project 1753), and the findings have been pre-viewed before publication to ensure that privacy sensitive, individual-specific information is not revealed. The data from this study can only be applied through a government data-sharing portal of Statistics Netherlands (https://www.cbs.nl/en-gb/onze-diensten/customised-services-microdata). Under certain conditions, these microdata are accessible for statistical and scientific research. For further information: microdata@cbs.nl.

References

Abrevaya, J. and C. M. Dahl (2008). The effects of birth inputs on birthweight: Evidence from quantile estimation on panel data. *Journal of Business & Economic Statistics 26*(4), 379–397.

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *The Quarterly Journal of Economics 97*(4), 543–569.

Akerlof, G. A. and J. L. Yellen (1988). Fairness and unemployment. *The American Economic Review 78*(2), 44–49.

Akerlof, G. A. and J. L. Yellen (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics 105*(2), 255–283.

Almond, D., K. Y. Chay, and D. S. Lee (2005). The costs of low birth weight. *The Quarterly Journal of Economics 120*(3), 1031–1083.

Almond, D. and B. Mazumder (2011). Health capital and the prenatal environment: The effect of Ramadan observance during pregnancy. *American Economic Journal: Applied Economics 3*(4), 56–85.

Altındağ, O., J. Greve, and E. Tekin (2024). Public health policy at scale: Impact of a government-sponsored information campaign on infant mortality in Denmark. *Review of Economics and Statistics 106*(3), 882–893.

Amelink-Verburg, M. P. and S. E. Buitendijk (2010). Pregnancy and labour in the Dutch maternity care system: What is normal? The role division between midwives and obstetricians. *Journal of Midwifery and Women's Health 55*(3), 216–225.

Bais, J. M. and M. Pel (2006). The basis of the Dutch obstetric system: Risk selection. *European Clinics in Obstetrics and Gynaecology 2*(4), 209–212.

Barreca, A. and M. Page (2015). A pint for a pound? Minimum drinking age laws and birth outcomes. *Health Economics 24*(4), 400–418.

Bekedam, D. J., S. Engelsbel, B. W. Mol, S. E. Buitendijk, K. M. van der Pal-de, et al. (2002). Male predominance in fetal distress during labor. *American Journal of Obstetrics and Gynecology 187*(6), 1605–1607.

Bertens, L. C., L. B. Ochoa, T. Van Ourti, E. A. Steegers, and J. V. Been (2020). Persisting inequalities in birth outcomes related to neighbourhood deprivation. *Journal of Epidemiology and Community Health 74*(3), 232–239.

Björkman Nyqvist, M., D. De Walque, and J. Svensson (2017). Experimental evidence on the long-run impact of community-based monitoring. *American Economic Journal: Applied Economics 9*(1), 33–69.

Blencowe, H., J. Krasevec, M. De Onis, R. E. Black, X. An, G. A. Stevens, E. Borghi, C. Hayashi, D. Estevez, L. Cegolon, et al. (2019). National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: A systematic analysis. *The Lancet Global Health 7*(7), e849–e860.

Bolten, N., A. De Jonge, E. Zwagerman, P. Zwagerman, T. Klomp, J. Zwart, and C. Geerts (2016). Effect of planned place of birth on obstetric interventions and maternal outcomes among low-risk women: A cohort study in the Netherlands. *BMC pregnancy and childbirth 16*, 1–13.

Broere-Brown, Z. A., E. Baan, S. Schalekamp-Timmermans, B. O. Verburg, V. W. Jaddoe, and E. A. Steegers (2016). Sex-specific differences in fetal and infant growth patterns: A prospective population-based cohort study. *Biology of sex differences 7*(1), 1–9.

Brosig-Koch, J., N. Kairies-Schwarz, and J. Kokot (2017). Sorting into payment schemes and medical treatment: A laboratory experiment. *Health Economics 26*, 52–65.

Buisman, P. and G. Gerats (2008). Verloskundigen in tijd erop vooruit in achterstandswijken. *Tijdschrift voor Verloskundigen 33*(9), 13.

Cabrera, J. M. and D. Webbink (2020). Do higher salaries yield better teachers and better student outcomes? *Journal of Human Resources 55*(4), 1222–1257.

Cadena, B. C. and A. C. Smith (2022). Performance pay, productivity, and strategic opt-out: Evidence from a community health center. *Journal of Public Economics 206*, 104580.

Calonico, S., M. D. Cattaneo, and M. H. Farrell (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal 23*(2), 192–210.

Centraal Bureau voor de Statistiek (2018). Kerncijfers wijken en buurten https://www.cbs.nl/nl-nl/reeksen/kerncijfers-wijken-en-buurten .

Collins, K. A. and E. Popek (2018). Birth injury: Birth asphyxia and birth trauma. *Academic Forensic Pathology 8*(4), 788–864.

Crayford, T., J. Shanks, M. Bajekal, S. Langford, and M. Bajekal (1995). Analysis from inner London of deprivation payments based on enumeration districts rather than wards. *British Medical Journal 311*(7008), 787–788.

Currie, J. (2011). Inequality at birth: Some causes and consequences. *American Economic Review 101*(3), 1–22.

Currie, J. and J. Gruber (1996). Health insurance eligibility, utilization of medical care, and child health. *The Quarterly Journal of Economics 111*(2), 431–466.

Cygan-Rehm, K. and K. Karbownik (2022). The effects of incentivizing early prenatal care on infant health. *Journal of Health Economics 83*, 102612.

Dahlen, H., M. Jackson, V. Schmied, S. Tracy, and H. Priddis (2011). Birth centres and the national maternity services review: Response to consumer demand or compromise? *Women and Birth 24*(4), 165–172.

Daysal, N. M., M. Trandafir, and R. Van Ewijk (2015). Saving lives at birth: The impact of home births on infant outcomes. *American Economic Journal: Applied Economics 7*(3), 28–50.

Daysal, N. M., M. Trandafir, and R. Van Ewijk (2019). Low-risk isn't no-risk: Perinatal treatments and the health of low-income newborns. *Journal of Health Economics 64*, 55–67.

De Onis, M. and J.-P. Habicht (1996). Anthropometric reference data for international use: recommendations from a World Health Organization Expert Committee. *The American Journal of Clinical Nutrition 64*(4), 650–658.

De Ree, J., K. Muralidharan, M. Pradhan, and H. Rogers (2018). Double for nothing? Experimental evidence on an unconditional teacher salary increase in Indonesia. *The Quarterly Journal of Economics 133*(2), 993–1039.

De Walque, D., E. Kandpal, A. Wagstaff, J. Friedman, M. Piatti-Fünfkirchen, A. Sautmann, G. Shapira, and E. Van de Poel (2022). *Improving effective coverage in health: Do financial incentives work?* World Bank Publications.

Devillé, W. and T. Wiegers (2012). Herijking stedelijke achterstandsgebieden 2012. Technical report, NIVEL.

Devlin, R. A. and S. Sarma (2008). Do physician remuneration schemes matter? The case of Canadian family physicians. *Journal of Health Economics 27*(5), 1168–1181.

Eijkenaar, F. (2012). Pay for performance in health care: An international overview of initiatives. *Medical Care Research and Review 69*(3), 251–276.

Eijkenaar, F., M. Emmert, M. Scheppach, and O. Schöffski (2013). Effects of pay for performance in health care: A systematic review of systematic reviews. *Health Policy 110*(2-3), 115–130.

Enthoven, A. C. and W. P. van de Ven (2007). Going Dutch: Managed-competition health insurance in the Netherlands. *New England Journal of Medicine 357*(24), 2421–2423.

Eskes, M., A. J. Waelput, S. A. Scherjon, K. A. Bergman, A. Abu-Hanna, and A. C. Ravelli (2017). Small for gestational age and perinatal mortality at term: An audit in a Dutch national cohort study. *European Journal of Obstetrics & Gynecology and Reproductive Biology 215*, 62–67.

Esteves-Sorenson, C. (2018). Gift exchange in the workplace: Addressing the conflicting evidence with a careful test. *Management Science 64*(9), 4365–4388.

Evers, A. C. C., H. A. A. Brouwers, C. W. P. M. Hukkelhoven, P. G. J. Nikkels, J. Boon, A. van Egmond-Linden, J. Hillegersberg, Y. S. Snuif, S. Sterken-Hooisma, H. W. Bruinse, and A. Kwee (2010). Perinatal mortality and severe morbidity in low and high risk term pregnancies in the Netherlands: Prospective cohort study. *British Medical Journal 341*, c5639.

Gluckman, P. D., W. Cutfield, P. Hofman, and M. A. Hanson (2005). The fetal, neonatal, and infant environments: The long-term consequences for disease risk. *Early Human Development 81*(1), 51–59.

Grembi, V., T. Nannicini, and U. Troiano (2016). Do fiscal rules matter? *American Economic Journal: Applied Economics 8*(3), 1–30.

Hahn, J., P. Todd, and W. Van der Klaauw (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica 69*(1), 201–209.

Hendrix, M. J., S. M. Evers, M. Basten, J. G. Nijhuis, and J. L. Severens (2009). Cost analysis of the Dutch obstetric system: Low-risk nulliparous women preferring home or short-stay hospital birth-a prospective non-randomised controlled study. *BMC Health Services Research 9*(1), 1–16.

Hermus, M. A., I. C. Boesveld, M. Hitzert, A. Franx, J. de Graaf, E. Steegers, T. A. Wiegers, and K. M. van der Pal-de Bruin (2017). Defining and describing birth centres in the Netherlands: a component study of the Dutch Birth Centre Study. *BMC Pregnancy and Childbirth 17*(1), 1–11.

Hingstman, L. and R. Kenens (2009). Cijfers uit de registratie van verloskundigen: Peiling 2008. Technical report, Nederlands Instituut voor Onderzoek van de Gezondheidszorg.

Hingstman, L. and R. Kenens (2010). Cijfers uit de registratie van verloskundigen: Peiling 2009. Technical report, Nederlands Instituut voor Onderzoek van de Gezondheidszorg.

Hingstman, L. and R. Kenens (2011). Cijfers uit de registratie van verloskundigen: Peiling 2010. Technical report, Nederlands Instituut voor Onderzoek van de Gezondheidszorg.

Hingstman, L. and R. Kenens (2012). Cijfers uit de registratie van verloskundigen: Peiling 2011. Technical report, Nederlands Instituut voor Onderzoek van de Gezondheidszorg.

Hobbs, R. (1993). Deprivation payments. *British Medical Journal 306*(6877), 534.

Hoftiezer, L., M. H. Hof, J. Dijs-Elsinga, M. Hogeveen, C. W. Hukkelhoven, and R. A. van Lingen (2019). From population reference to national standard: New and improved birthweight charts. *American Journal of Obstetrics and Gynecology 220*(4), 383–e1.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental analysis of neighborhood effects. *Econometrica 75*(1), 83–119.

Kraemer, S. (2000). The fragile male. *British Medical Journal 321*(7276), 1609–1612.

Laureij, L. T., M. Van Der Hulst, J. Lagendijk, J. V. Been, H. E. Ernst-Smelt, A. Franx, and M. Lugtenberg (2021). Insight into the process of postpartum care utilisation and in-home support among vulnerable women in the Netherlands: An in-depth qualitative exploration. *British Medical Journal Open 11*(9), e046696.

Lavoie, M. E., P. Robaey, J. E. Stauder, J. Glorieux, and F. Lefebvre (1998). Extreme prematurity in healthy 5-year-old children: A re-analysis of sex effects on event-related brain activity. *Psychophysiology 35*(6), 679–689.

Lazuka, V. (2018). The long-term health benefits of receiving treatment from qualified midwives at birth. *Journal of Development Economics 133*, 415–433.

Lee, A. C., J. Katz, H. Blencowe, S. Cousens, N. Kozuki, J. P. Vogel, L. Adair, A. H. Baqui, Z. A. Bhutta, L. E. Caulfield, et al. (2013). National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010. *The Lancet Global Health 1*(1), e26–e36.

Lescure, D., S. Schepman, R. Batenburg, T. A. Wiegers, and E. Verbakel (2017). Preferences for birth center care in the Netherlands: An exploration of ethnic differences. *BMC Pregnancy and Childbirth 17*(1), 1–10.

Maas, V. Y., M. Poels, M. H. de Kievit, A. P. Hartog, A. Franx, and M. P. Koster (2022). Planning is not equivalent to preparing, how Dutch women perceive their pregnancy planning in relation to preconceptional lifestyle behaviour change: A cross-sectional study. *BMC Pregnancy and Childbirth 22*(1), 577.

Manniën, J., T. Klomp, T. Wiegers, M. Pereboom, J. Brug, A. de Jonge, M. van der Meijde, E. Hutton, F. Schellevis, and E. Spelten (2012). Evaluation of primary care midwifery in the Netherlands: Design and rationale of a dynamic cohort study (DELIVER). *BMC Health Services Research 12*, 1–11.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

Mizuno, R. (2000). The male/female ratio of fetal deaths and births in Japan. *The Lancet 356*(9231), 738–739.

Morris, R. K., E. Johnstone, C. Lees, V. Morton, G. Smith, R. C. of Obstetricians, and Gynaecologists (2024). Investigation and care of a small-for-gestational-age fetus and a growth restricted fetus (green-top guideline no. 31). *BJOG: An International Journal of Obstetrics & Gynaecology 131*(9), e31–e80.

Naeye, R. L., L. S. Burt, D. L. Wright, W. A. Blanc, and D. Tatter (1971). Neonatal mortality, the male disadvantage. *Pediatrics 48*(6), 902–906.

National Health Care Institute (2006). The Dutch health care system.

National Health Service (2019). Caesarean section - risks. https://www.nhs.uk/conditions/caesarean-section/risks/. Accessed January 2025.

Neal, D. (2011). The design of performance pay in education. In *Handbook of the Economics of Education*, Volume 4, pp. 495–550. Elsevier.

Nederlandse Zorgautoriteit (2008). NZA jaarsverlag 2008.

Nederlandse Zorgautoriteit (2009). Verloskundigen met bijlagen - 5700-1900-09-1-volgnr.21.

Nilsson, J. P. (2017). Alcohol availability, prenatal conditions, and long-term economic outcomes. *Journal of Political Economy 125*(4), 1149–1207.

Organisation for Economic Co-operation and Development (2008). Equivalence scales.

Perry, H. B., R. Zulliger, and M. M. Rogers (2014). Community health workers in low-, middle-, and high-income countries: An overview of their history, recent evolution, and current effectiveness. *Annual Review of Public Health 35*(1), 399–421.

Petersen, L. A., L. D. Woodard, T. Urech, C. Daw, and S. Sookanan (2006). Does pay-for-performance improve the quality of health care? *Annals of Internal Medicine 145*(4), 265–272.

Rosenzweig, M. R. and T. P. Schultz (1983). Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *Journal of political economy 91*(5), 723–746.

Rowe, R. E., J. J. Kurinczuk, L. Locock, and R. Fitzpatrick (2012). Women's experience of transfer from midwifery unit to hospital obstetric unit during labour: A qualitative interview study. *BMC Pregnancy and Childbirth 12*, 1–15.

Sanders, N. J. and C. Stoecker (2015). Where have all the young men gone? Using sex ratios to measure fetal death rates. *Journal of Health Economics 41*, 30–45.

Schäfer, W., M. Kroneman, W. Boerma, M. Van den Berg, G. Westert, W. Devillé, E. Van Ginneken, W. H. Organization, et al. (2010). The Netherlands: Health system review. Technical report, World Health Organization.

Schlaudecker, E. P., F. M. Munoz, A. Bardají, N. S. Boghossian, A. Khalil, H. Mousa, M. Nesin, M. I. Nisar, V. Pool, H. M. Spiegel, et al. (2017). Small for gestational age: Case definition & guidelines for data collection, analysis, and presentation of maternal immunisation safety data. *Vaccine 35*(48Part A), 6518.

Shaikh, E. M., S. Mehmood, and M. A. Shaikh (2010). Neonatal outcome in meconium stained amniotic fluid-one year experience. *Journal of the Pakistan Medical Association 60*(9), 711–4.

Shapiro, C. and J. E. Stiglitz (1984). Equilibrium unemployment as a worker discipline device. *The American Economic Review 74*(3), 433–444.

Sheldon, T. (2008). Perinatal mortality in Netherlands third worst in Europe. *British Medical Journal 337*, a3118.

Singer, J. E., M. Westphal, and K. R. Niswander (1968). Sex differences in the incidence of neonatal abnormalities and abnormal performance in early childhood. *Child Development 39*(1), 103–112.

Sonchak, L. (2015). Medicaid reimbursement, prenatal care and infant health. *Journal of Health Economics 44*, 10–24.

Statistics Netherlands (2024). Person with a western migration background. https://www.cbs.nl/en-gb/our-services/methods/definitions/person-with-a-western-migration-background. Accessed: 2024-10-11.

Van de Ven, W. P. and F. T. Schut (2008). Universal mandatory health insurance in the Netherlands: A model for the United States? *Health Affairs 27*(3), 771–781.

van der Geest, S. A. and M. Varkevisser (2019). Patient responsiveness to a differential deductible: Empirical results from The Netherlands. *The European Journal of Health Economics 20*, 513–524.

van Teijlingen, E. and P. McCaffery (1987). The profession of midwife in the Netherlands. *Midwifery 3*(4), 178–186.

Verheij, R. A., D. H. De Bakker, and S. A. Reijneveld (2001). GP income in relation to workload in deprived urban areas in the Netherlands: Before and after the 1996 pay review. *The European Journal of Public Health 11*(3), 264–266.

Vidiella-Martin, J., J. V. Been, E. Van Doorslaer, P. García-Gómez, and T. Van Ourti (2021). Association between income and perinatal mortality in the Netherlands across gestational age. *JAMA Network Open 4*(11), e2132124–e2132124.

Vos, A. A., A. G. Posthumus, G. J. Bonsel, E. A. Steegers, and S. Denktaş (2014). Deprived neighborhoods and adverse perinatal outcome: A systematic review and meta-analysis. *Acta Obstetricia et Gynecologica Scandinavica 93*(8), 727–740.

Vos, A. A., S. F. van Voorst, E. A. Steegers, and S. Denktaş (2016). Analysis of policy towards improvement of perinatal mortality in the Netherlands (2004–2011). *Social Science & Medicine 157*, 156–164.

Vries, R. D., T. A. Wiegers, B. Smulders, and E. V. Teijlingen (2012). The Dutch Obstetrical System: Vanguard of the Future in Maternity Care. In *Birth Models That Work*. University of California Press Berkeley.

Wiegerinck, M. M., M. Eskes, J. A. van der Post, B. W. Mol, and A. C. Ravelli (2020). Intrapartum and neonatal mortality in low-risk term women in midwife-led care and obstetrician-led care at the onset of labor: A national matched cohort study. *Acta Obstetricia et Gynecologica Scandinavica 99*(4), 546–554.

Wiegers, T. and W. Devillé (2008). Herijking stedelijke achterstandsgebieden 2008. Technical report, NIVEL.

Wiegers, T. A. and C. W. Hukkelhoven (2010). The role of hospital midwives in the Netherlands. *BMC Pregnancy and Childbirth 10*(1), 1–9.

World Health Organization (1995). Physical Status: The Use and Interpretation of Anthropometry: Report of a WHO Expert Committee.

Worrall, A., J. N. Rea, and Y. Ben-Shlomo (1997). Counting the cost of social disadvantage in primary care: Retrospective analysis of patient data. *British Medical Journal 314*(7073), 38.

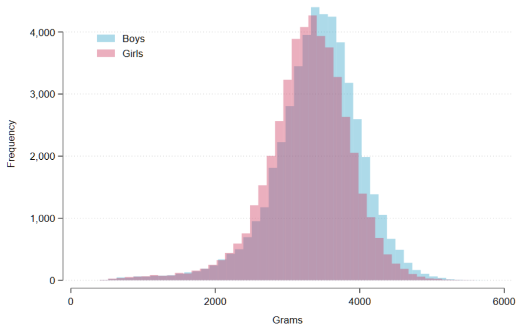
Yoder, B. A., E. A. Kirsch, W. H. Barth Jr, and M. C. Gordon (2002). Changing obstetric practices associated with decreasing incidence of meconium aspiration syndrome. *Obstetrics & Gynecology 99*(5), 731–739.

Zeitlin, J., A. Mohangoo, M. Delnord, M. Cuttini, E.-P. S. Committee, et al. (2013). The second European Perinatal Health Report: Documenting changes over 6 years in the health of mothers and babies in Europe. *Journal of Epidemiology and Community Health 67*(12), 983–985.

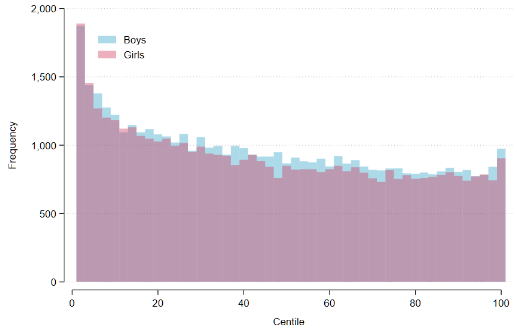
Figures

Figure 1: Distribution of Continuous Outcomes

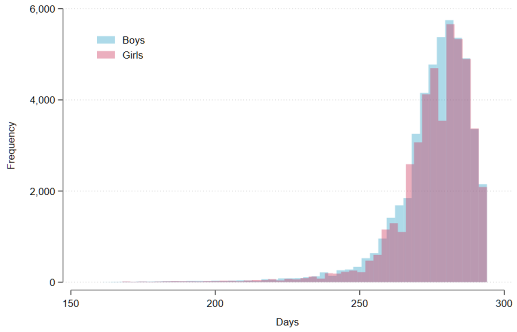
(a) Birthweight



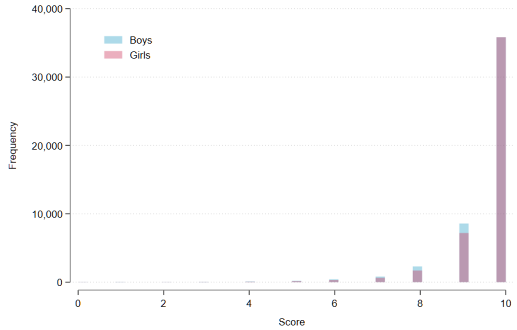
(b) Birthweight centile



(c) Gestational age



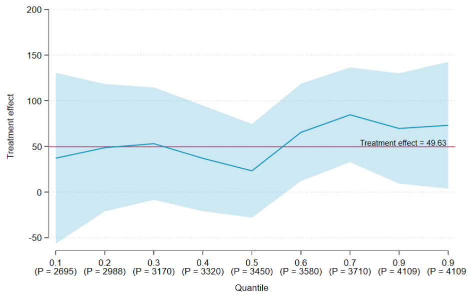
(d) Apgar 5



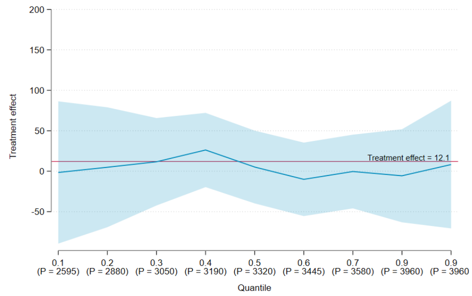
Notes: This figure displays the distribution of key continuous health outcomes at birth for boys and girls separately. Panel (a) shows birthweight in grams, panel (b) presents birthweight centile (adjusted for gestational age and sex), panel (c) shows gestational age in days, and panel (d) presents the Apgar score at 5 minutes (ranging from 0 to 10). Histograms are overlaid for boys (blue) and girls (red).

Figure 2: Quantile Regressions

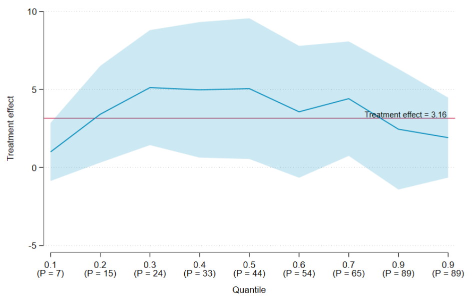
(a) Boys: birthweight



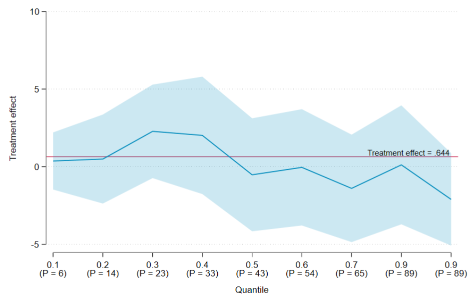
(b) Girls: birthweight



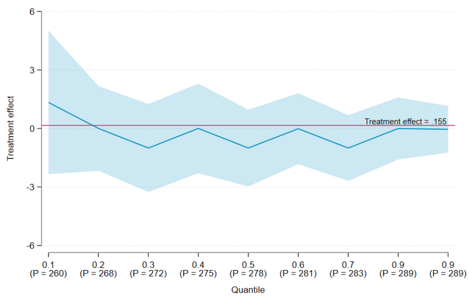
(c) Boys: birthweight centile



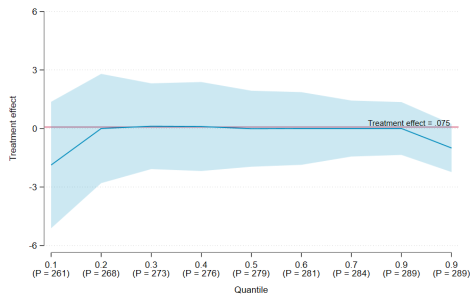
(d) Girls: birthweight centile



(e) Boys: gestational age



(f) Girls: gestational age



Notes: This figure presents point estimates from quantile regressions on continuous birth outcomes, ranging from the 10^{th} to the 90^{th} percentile of the outcome distributions. The blue lines represent the estimated effects at each quantile, while the shaded areas denote the 95% confidence intervals. Standard errors are clustered at the neighborhood level. The red lines indicate the average treatment effects reported in Table A.4. Estimates for Apgar scores and the summary index are omitted because percentiles 20 to 100 consistently show an Apgar score of 10.

Tables

Table 1: Descriptive Statistics

	Study population				All births			
	Boys		Girls		Boys		Girls	
	(1)		(2)		(3)		(4)	
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
Panel A: Birth Outcomes								
Birthweight (grams)	48,407	3,406.9	46,018	3,283.9	324,012	3,482.1	307,476	3,355.6
		(605.6)		(582.6)		(609.1)		(579.9)
Birthweight centile (1-100)	48,407	45.8	46,018	45.4	319,298	50.1	302,946	49.6
		(29.6)		(29.8)		(29.8)		(29.8)
Gestational age (days)	48,407	275.4	46,018	276.0	323,255	275.9	306,648	276.5
		(14.1)		(13.7)		(14.0)		(13.5)
Apgar score (1-10)	48,407	9.6	46,018	9.7	323,921	9.6	307,386	9.7
		(1.0)		(0.9)		(0.9)		(0.8)
Early neonatal mortality (%)	48,407	0.2	46,018	0.2	324,166	0.2	307,610	0.2
Low birthweight (%)	48,407	6.5	46,018	7.7	324,012	5.6	307,476	6.6
Small for gestational age (%)	48,407	14.8	46,018	15.2	319,298	11.9	302,946	12.1
Preterm (%)	48,407	8.1	46,018	7.3	323,255	7.7	306,648	6.9
Apgar <7 (%)	48,407	1.9	46,018	1.4	323,921	1.5	307,386	1.2
Panel B: Risk Profiles, Midwifery Care and Fetal Distress								
Low risk (%)	48,407	52.9	46,018	52.9	322,020	53.3	305,587	53.7
Birth ends supervised by a community midwife (%)	48,407	25.5	46,018	27.3	323,550	28.1	306,968	29.9
Emergency Caesarean section (%)	48,407	10.7	46,018	8.6	324,166	10.2	307,610	8.5
Meconium in the amniotic fluid (%)	48,407	7.2	46,018	7.8	324,166	6.0	307,610	6.3
Birth trauma/perinatal asphyxia (%)	48,407	0.6	46,018	0.4	324,166	0.6	307,610	0.4
Panel C: Demographic and Socioeconomic Characteristics								
Primiparous (%)	48,407	47.8	46,016	47.5	324,152	46.4	307,592	46.2
Multiple birth (%)	48,407	1.5	46,018	1.6	324,166	1.7	307,160	1.7
Birth order	48,407	2.4	46,018	2.4	324,166	2.3	307,610	2.3
		(1.1)		(1.1)		(1.0)		(1.0)
Maternal age	48,407	29.6	46,018	29.5	324,164	30.5	307,609	30.5
		(5.4)		(5.4)		(5.0)		(5.0)
Dutch mother (%)	48,407	59.4	46,017	59.6	324,166	80.8	307,609	81.1
Paternal age	43,683	33.8	41,419	33.8	311,020	34.1	294,795	34.1
		(6.6)		(6.6)		(5.8)		(5.8)
Dutch father (%)	43,682	57.0	41,419	57.2	311,019	82.0	294,795	82.1
Standardized income (euro)	46,828	20,393.0	44,458	20,316.0	317,165	24,782.9	300,965	24,795.4
		(11,030.7)		(10,680.2)		(13,745.9)		(13,672.0)
N	48,407		46,018		324,166		307,610	

Notes: This table reports summary statistics for the study population (columns 1 and 2) and all births in the sample period (columns 3 and 4). For each variable, the number of available observations is reported first, followed by the mean and standard deviation (SD) in parentheses. Panel A includes key health outcomes at birth: birth weight, birth weight centile (adjusted for gestational age and sex), gestational age, Apgar score (1-10), early neonatal mortality, and binary indicators for adverse outcomes such as low birth weight, small for gestational age, preterm birth, and Apgar score <7. Panel B includes indicators for risk profiles and birth-related characteristics, including whether the pregnancy was classified as low risk, whether the birth occurred at home, and whether it was supervised by a community midwife. Measures of fetal distress and delivery outcomes are also included, such as emergency Caesarean section and birth trauma. Panel C provides demographic and socioeconomic characteristics, including maternal age, paternal age, Dutch nationality of the parents, primiparity, multiple births, and standardized household income (in euros). Disposable income consists of total household income, including income transfers, but net of income taxes and contributionssuch as social security contributions and health insurance premiums. The OECD modified equivalence scale (Organisation for Economic Co-operation and Development, 2008) is used to standardize for the household size in which the mother is registered. Income and household size refer to the year before delivery, unaffected by the survival of a newborn.

Table 2: Effect of Deprivation Payment on Birth Outcomes

	(1)	(2)	(3)	(4)	(5)
	Low birthweight	Small for gestational age	Preterm	Apgar <7	Summary index
Treatment effect	-0.0174*	-0.0322**	-0.00831	-0.0120**	0.108***
	(0.0103)	(0.0150)	(0.0115)	(0.00502)	(0.0405)
Mean DV	0.065	0.148	0.081	0.019	-0.000
N	48,407	48,407	48,407	48,407	48,407

(a) Boys

	(1)	(2)	(3)	(4)	(5)
	Low birthweight	Small for gestational age	Preterm	Apgar <7	Summary index
Treatment effect	0.00984	-0.0197	0.00485	0.00544	-0.0179
	(0.0118)	(0.0143)	(0.0101)	(0.00516)	(0.0407)
Mean DV	0.077	0.152	0.073	0.014	0.000
N	46,018	46,018	46,018	46,018	46,018

(b) Girls

Notes: This table presents estimates of the effect of deprivation payments on key birth outcomes. The dependent variables in columns (1)-(4) are binary indicators for: low birth weight (birth weight <2,500g), small for gestational age (birth weight below the 10th percentile, adjusted for gestational age and sex), preterm birth (gestational age <37 weeks), and low Apgar score (<7 at 5 minutes). Column (5) reports estimates for a summary index aggregating these outcomes into a single measure, standardized to have a mean of 0 and a standard deviation of 1. Following Kling et al. (2007), the summary index is constructed as the sum of z-scores of its components, with each component oriented so that higher scores indicate more beneficial outcomes. A positive treatment effect on the summary index thus reflects improved health at birth. The estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * p< 0.1, ** p< 0.05, *** p< 0.01.

Table 3: Effect of Deprivation Payment on Birth Outcomes, by Risk Profiles

	(1) Low birthweight	(2) Small for gestational age	(3) Preterm	(4) Apgar <7	(5) Summary index
Treatment effect	-0.0108	-0.0210	-0.00832	-0.00771	0.118**
	(0.00724)	(0.0179)	(0.00808)	(0.00509)	(0.0518)
Mean DV	0.017	0.119	0.026	0.010	0.000
N	25,605	25,605	25,605	25,605	25,605

(a) Boys: low risk

	(1) Low birthweight	(2) Small for gestational age	(3) Preterm	(4) Apgar <7	(5) Summary index
Treatment effect	-0.0282	-0.0472**	-0.00986	-0.0174*	0.126**
	(0.0202)	(0.0239)	(0.0210)	(0.00981)	(0.0589)
Mean DV	0.119	0.182	0.143	0.028	-0.000
N	22,802	22,802	22,802	22,802	22,802

(b) Boys: high risk

	(1) Low birthweight	(2) Small for gestational age	(3) Preterm	(4) Apgar <7	(5) Summary index
Treatment effect	0.00343	0.00130	0.00563	0.00190	-0.0339
	(0.00712)	(0.0183)	(0.00852)	(0.00479)	(0.0504)
Mean DV	0.020	0.114	0.022	0.008	-0.000
N	24,354	24,354	24,354	24,354	24,354

(c) Girls: low risk

	(1) Low birthweight	(2) Small for gestational age	(3) Preterm	(4) Apgar <7	(5) Summary index
Treatment effect	0.0166	-0.0456*	0.00298	0.00963	-0.00850
	(0.0226)	(0.0241)	(0.0191)	(0.00873)	(0.0603)
Mean DV	0.141	0.195	0.130	0.021	-0.000
N	21,664	21,664	21,664	21,664	21,664

(d) Girls: high risk

Notes: This table presents estimates of the effect of deprivation payments on key birth outcomes, separately by risk profile and sex. The dependent variables in columns (1)-(4) are binary indicators for: low birth weight (birth weight <2,500g), small for gestational age (birth weight below the 10th percentile, adjusted for gestational age and sex), preterm birth (gestational age <37 weeks), and low Apgar score (<7 at 5 minutes). Column (5) reports estimates for a summary index aggregating these outcomes into a single measure, standardized to have a mean of 0 and a standard deviation of 1. Following Kling et al. (2007), the summary index is constructed as the sum of z-scores of its components, with each component oriented so that higher scores indicate more beneficial outcomes. A positive treatment effect on the summary index thus reflects improved health at birth. Pregnancies are categorized as low-risk if delivery begins under the supervision of community midwives (first tier of care) and as high-risk if delivery starts under the supervision of clinical midwives or OB/GYNs (secondary tier of care). The estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * p< 0.1, ** p< 0.05, *** p< 0.01.

Table 4: Effect of Deprivation Payment on Risk Selection Procedures by Midwives

	Boys	Girls
	(1)	(2)
	Low risk	
Treatment effect	-0.00576	0.00504
	(0.0224)	(0.0229)
Mean DV	0.529	0.529
N	48,407	46,018

Notes: This table presents estimates of the effect of deprivation payments on pregnancy risk classification, measured by the likelihood of being categorized as *Low risk*. The dependent variable, *Low risk*, equals 1 if delivery began under the supervision of community midwives (first tier of care) and 0 otherwise. Estimates are reported separately for male and female pregnancies. Pregnancies are categorized as low-risk if delivery begins under the supervision of community midwives, and as high-risk if delivery starts under the supervision of clinical midwives or OB/GYNs (secondary/tertiary tiers of care). The estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Effect of Deprivation Payment on Midwifery Outcomes, by Risk Profiles

	(1)	(2)	(3)	(4)	(5)
	Midwife ends	Emergency C-section	Meconium	Trauma	Summary index
Treatment effect	0.0167	-0.00605	-0.0352**	-0.000925	0.0861
	(0.0330)	(0.0146)	(0.0178)	(0.00359)	(0.0643)
Mean DV	0.483	0.072	0.116	0.005	0.000
N	25,605	25,605	25,605	25,605	25,605

(a) Boys: low risk

	(1)	(2)	(3)	(4)	(5)
	Midwife ends	Emergency C-section	Meconium	Trauma	Summary index
Treatment effect	0	0.00182	-0.00138	-0.00453	0.00306
	(.)	(0.0189)	(0.0118)	(0.00382)	(0.0621)
Mean DV	0.000	0.146	0.024	0.006	-0.000
N	22,802	22,802	22,802	22,802	22,802

(b) Boys: high risk

	(1)	(2)	(3)	(4)	(5)
	Midwife ends	Emergency C-section	Meconium	Trauma	Summary index
Treatment effect	-0.0396	-0.00958	0.0175	0.000989	-0.0546
	(0.0279)	(0.0126)	(0.0179)	(0.00331)	(0.0555)
Mean DV	0.515	0.053	0.124	0.003	-0.000
N	24,354	24,354	24,354	24,354	24,354

(c) Girls: low risk

	(1)	(2)	(3)	(4)	(5)
	Midwife ends	Emergency C-section	Meconium	Trauma	Summary index
Treatment effect	0	0.0207	-0.0101	0.00168	-0.0282
	(.)	(0.0222)	(0.0113)	(0.00369)	(0.0653)
Mean DV	0.000	0.124	0.026	0.004	0.000
N	21,664	21,664	21,664	21,664	21,664

(d) Girls: high risk

Notes: Each column represents a separate estimate from equation 1 on intermediate outcomes measuring midwifery performance, separately for boys and girls and high- and low-risk profiles. The four outcomes are binary indicators for: (1) whether the community midwife remains the responsible caregiver until the end of the delivery (low-risk pregnancies only), (2) the occurrence of emergency C-sections, (3) the presence of meconium in the amniotic fluid, and (4) the presence of birth trauma or perinatal asphyxia. Column (5) reports estimates for a summary index aggregating these outcomes into a single measure, standardized to have a mean of 0 and a standard deviation of 1. Following Kling et al. (2007), the summary index is constructed as the sum of z-scores of its components, with each component oriented so that higher scores indicate more beneficial outcomes. A positive treatment effect on the summary index thus reflects improved health at birth. Pregnancies are categorized as low-risk if delivery begins under the supervision of community midwives (first tier of care) and as high-risk if delivery starts under the supervision of clinical midwives or OB/GYNs (secondary tier of care). High-risk pregnancies, by definition, remain under the care of the secondary tier of care until delivery and are excluded from the first outcome. Standard errors are clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * p< 0.1, ** p< 0.05, *** p< 0.01.

Table 6: Effect of Deprivation Payment on Birth Outcomes Over Time

	(1) Low birthweight	(2) Small for gestational age	(3) Preterm	(4) Apgar <7	(5) Summary index
Treatment x 2009	-0.0287**	-0.0381**	-0.0189	-0.0120**	0.147***
	(0.0116)	(0.0184)	(0.0138)	(0.00570)	(0.0464)
Treatment x 2010	-0.00834	-0.0211	-0.0109	-0.0146**	0.0933*
	(0.0122)	(0.0186)	(0.0136)	(0.00656)	(0.0517)
Treatment x 2011	-0.0153	-0.0378**	0.00470	-0.00935	0.0849*
	(0.0130)	(0.0157)	(0.0140)	(0.00661)	(0.0493)
Mean DV	0.065	0.148	0.081	0.019	-0.000
N	48,407	48,407	48,407	48,407	48,407
F-test	0.200	0.473	0.235	0.736	0.381

(a) Boys

	(1) Low birthweight	(2) Small for gestational age	(3) Preterm	(4) Apgar <7	(5) Summary index
Treatment x 2009	0.0129	-0.00738	-0.00290	0.00497	-0.0225
	(0.0129)	(0.0185)	(0.0126)	(0.00591)	(0.0478)
Treatment x 2010	-0.00460	-0.0441**	0.000294	0.00826	0.0266
	(0.0174)	(0.0174)	(0.0144)	(0.00658)	(0.0609)
Treatment x 2011	0.0208	-0.00832	0.0165	0.00315	-0.0556
	(0.0133)	(0.0180)	(0.0123)	(0.00597)	(0.0486)
Mean DV	0.077	0.152	0.073	0.014	0.000
N	46,018	46,018	46,018	46,018	46,018
F-test	0.210	0.088	0.296	0.693	0.351

(b) Girls

Notes: This table presents estimates of the effect of deprivation payments on key birth outcomes, separately for boys and girls and disaggregated by post-intervention year (20092011). The dependent variables in columns (1)(4) are binary indicators for: low birth weight (birth weight <2,500g), small for gestational age (birth weight below the 10th percentile, adjusted for gestational age and sex), preterm birth (gestational age <37 weeks), and low Apgar score (<7 at 5 minutes). Column (5) reports estimates for a summary index aggregating these outcomes into a single measure, standardized to have a mean of 0 and a standard deviation of 1. Following Kling et al. (2007), the summary index is constructed as the sum of z-scores of its components, with each component oriented so that higher scores indicate more beneficial outcomes. A positive treatment effect on the summary index thus reflects improved health at birth. The estimates are based on a modified version of equation 1, which includes separate treatment effects for each year after the policys implementation. The *F-test* reports the p-value for the joint significance test of the coefficients *Treatment x 2009*, *Treatment x 2010*, and *Treatment x 2011*. Standard errors are clustered at the neighborhood level, which is the unit of treatment, and are reported in parentheses. Statistical significance is denoted by * p< 0.1, ** p< 0.05, *** p< 0.01.

Table 7: Effect of Deprivation Payment on Household Location

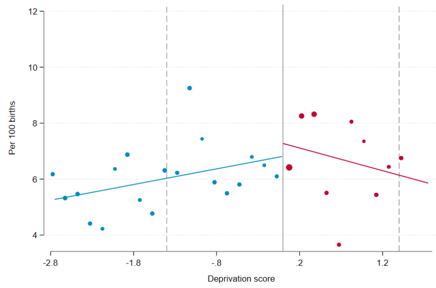
	Boys	Girls
	(1)	(2)
	Movers	
Treatment effect	-0.00493	-0.00422
	(0.0104)	(0.0120)
Mean DV	0.090	0.089
N	48,315	45,917

Notes: This table presents estimates of the effect of deprivation payments on maternal mobility, measured by the likelihood of mothers changing their residential location between six months before conception and delivery. The dependent variable, *Movers*, is a binary indicator equal to 1 if the mother changed her location and 0 otherwise. Estimates are reported separately for boys and girls. The number of observations differs from the baseline analyses because the residential location before delivery could not be determined for 92 boys and 101 girls. The estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * p< 0.1, ** p< 0.05, *** p< 0.01.

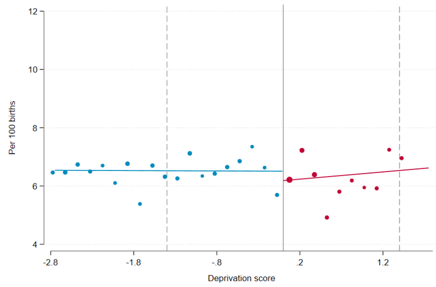
Appendix A: Additional Figures and Tables

Figure A.1: Birth Outcomes by Deprivation Score: Boys

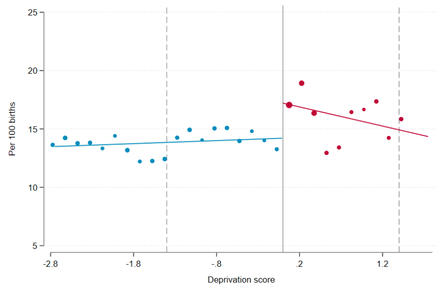
(a) Low birthweight: pre-intervention



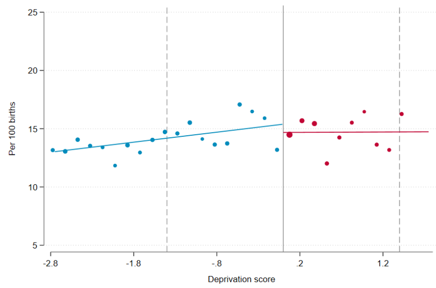
(b) Low birthweight: post-intervention



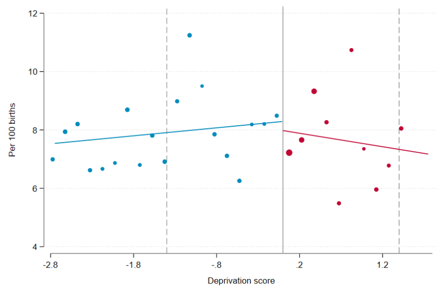
(c) SGA pre-intervention



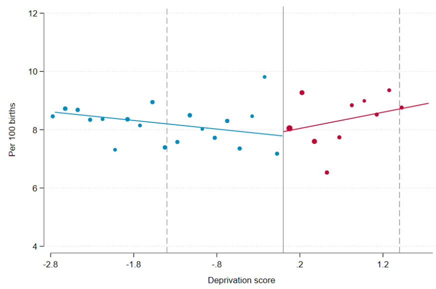
(d) SGA post-intervention



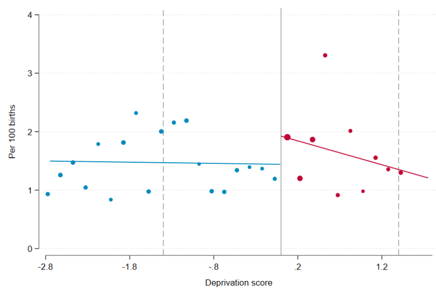
(e) Preterm: pre-intervention



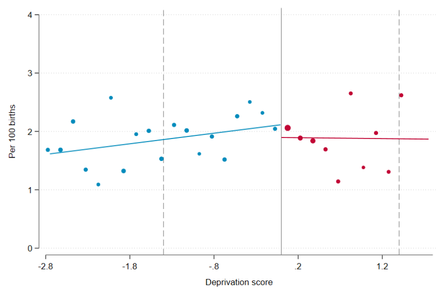
(f) Preterm: post-intervention



(g) Apgar <7: pre-intervention



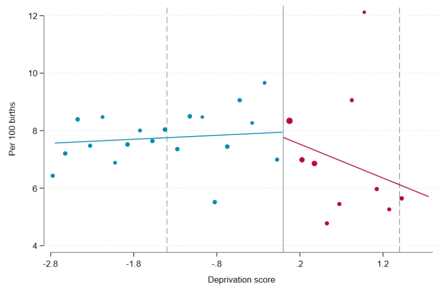
(h) Apgar <7: post-intervention



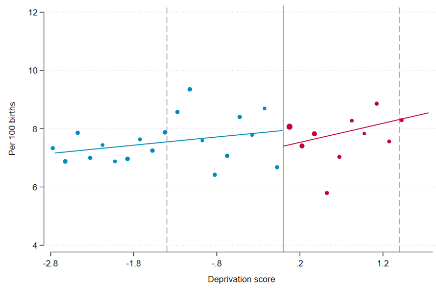
Notes: SGA = small for gestational age. Each dot represents the average outcome for a given deprivation score, centered at 0, and grouped into bins of 0.15 deprivation score units. The size of each marker is proportional to the number of observations in the bin. Observations within a bandwidth of $h = 2.8$ are used to fit the linear trends, which are estimated using individual-level data on either side of the threshold. Blue dots represent individuals in non-deprived neighborhoods, and red dots represent those in deprived neighborhoods. Panels (a), (c), (e), and (g) show outcomes from the pre-intervention period (2008), while panels (b), (d), (f), and (h) show outcomes from the post-intervention period (2009-2011). The vertical lines at $h = 1.4$ indicate the optimal bandwidth (see Section 4).

Figure A.2: Birth Outcomes by Deprivation Score: Girls

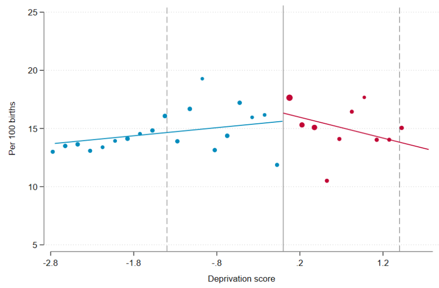
(a) Low birthweight: pre-intervention



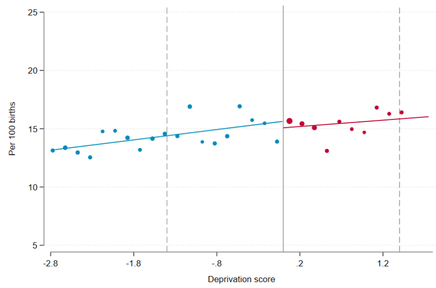
(b) Low birthweight: post-intervention



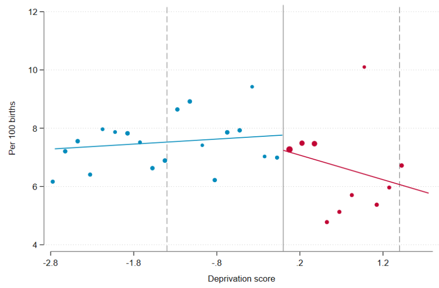
(c) SGA pre-intervention



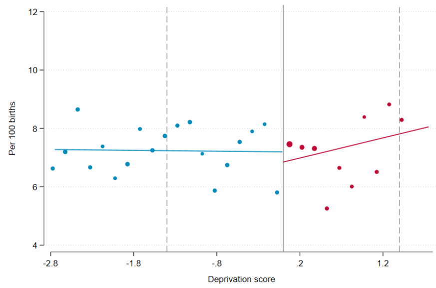
(d) SGA post-intervention



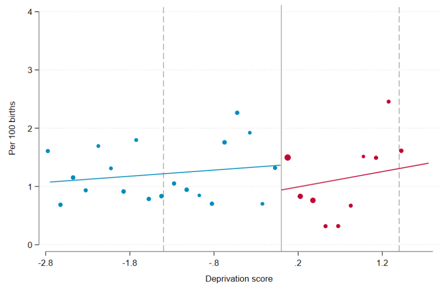
(e) Preterm: pre-intervention



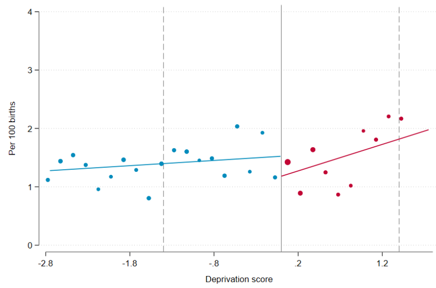
(f) Preterm: post-intervention



(g) Apgar <7: pre-intervention



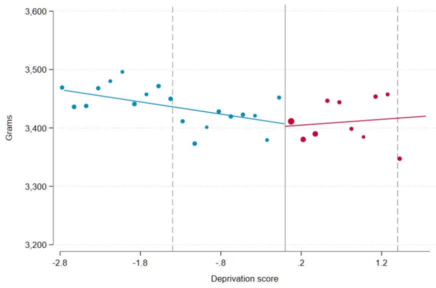
(h) Apgar <7: post-intervention



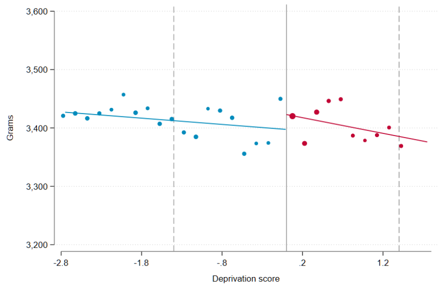
Notes: SGA = small for gestational age. Each dot represents the average outcome for a given deprivation score, centered at 0, and grouped into bins of 0.15 deprivation score units. The size of each marker is proportional to the number of observations in the bin. Observations within a bandwidth of $h = 2.8$ are used to fit the linear trends, which are estimated using individual-level data on either side of the threshold. Blue dots represent individuals in non-deprived neighborhoods, and red dots represent those in deprived neighborhoods. Panels (a), (c), (e), and (g) show outcomes from the pre-intervention period (2008), while panels (b), (d), (f), and (h) show outcomes from the post-intervention period (2009-2011). The vertical lines at $h = 1.4$ indicate the optimal bandwidth (see Section 4).

Figure A.3: Continuous Birth Outcomes by Deprivation Score Bins: Boys

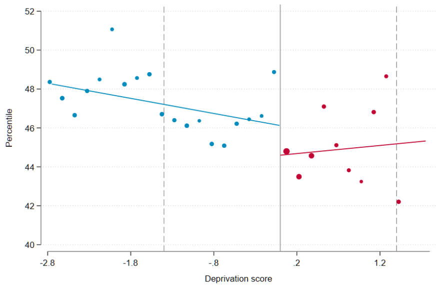
(a) Birthweight: pre-intervention



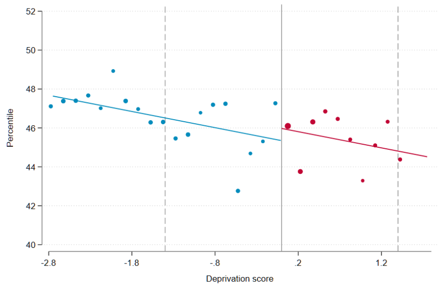
(b) Birthweight: post-intervention



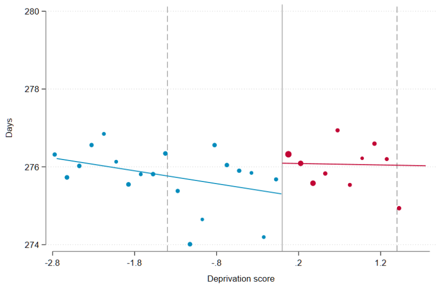
(c) Birthweight centile: pre-intervention



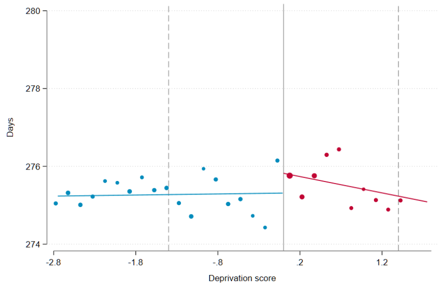
(d) Birthweight centile: post-intervention



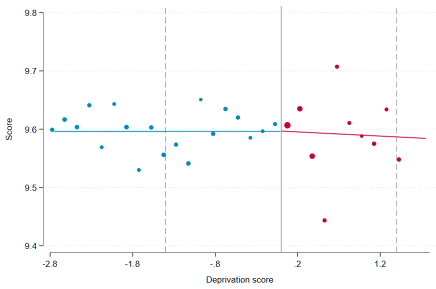
(e) Gestational age: pre-intervention



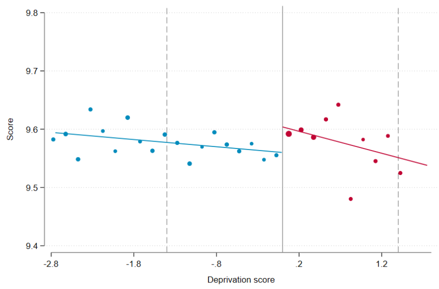
(f) Gestational age: post-intervention



(g) Apgar score: pre-intervention



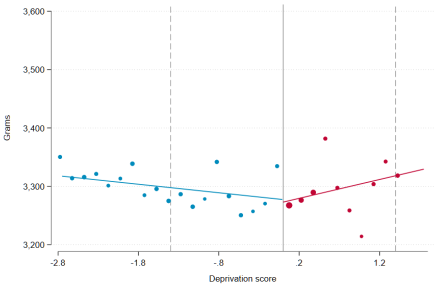
(h) Apgar score: post-intervention



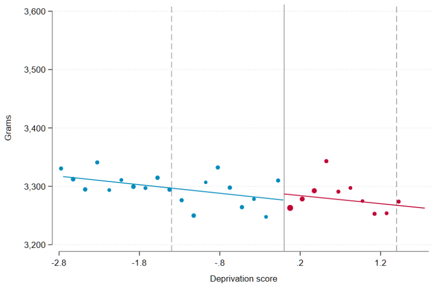
Notes: Each dot represents the average outcome for a given deprivation score, centered at 0, and grouped into bins of 0.15 deprivation score units. The size of each marker is proportional to the number of observations in the bin. Observations within a bandwidth of $h = 2.8$ are used to fit the linear trends, which are estimated using individual-level data on either side of the threshold. Blue dots represent individuals in non-deprived neighborhoods, and red dots represent those in deprived neighborhoods. Panels (a), (c), (e), and (g) show outcomes from the pre-intervention period (2008), while panels (b), (d), (f), and (h) show outcomes from the post-intervention period (2009-2011). The vertical lines at $h = 1.4$ indicate the optimal bandwidth (see Section 4).

Figure A.4: Continuous Birth Outcomes by Deprivation Score Bins: Girls

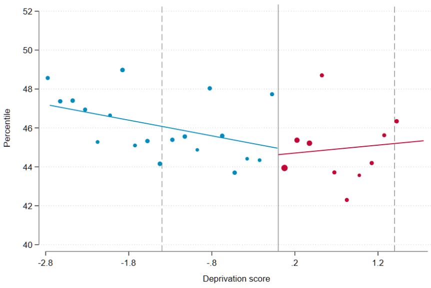
(a) Birthweight: pre-intervention



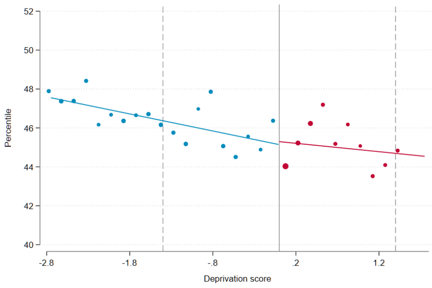
(b) Birthweight: post-intervention



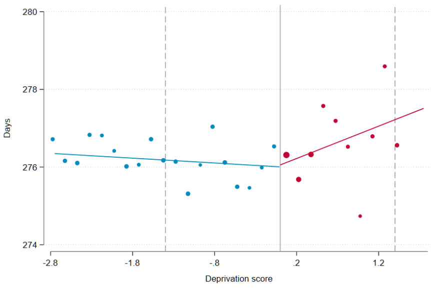
(c) Birthweight centile: pre-intervention



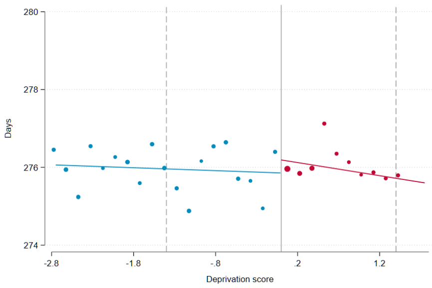
(d) Birthweight centile: post-intervention



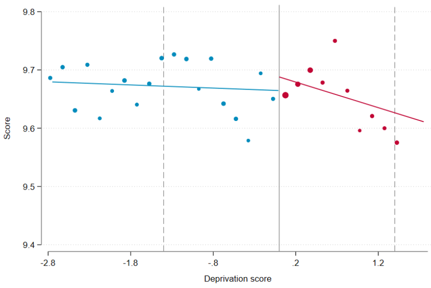
(e) Gestational age: pre-intervention



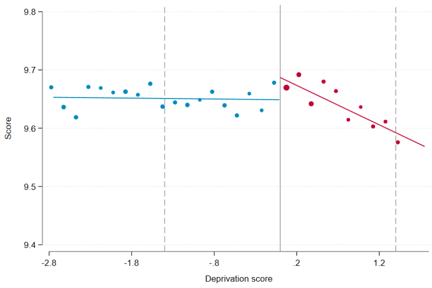
(f) Gestational age: post-intervention



(g) Apgar score: pre-intervention

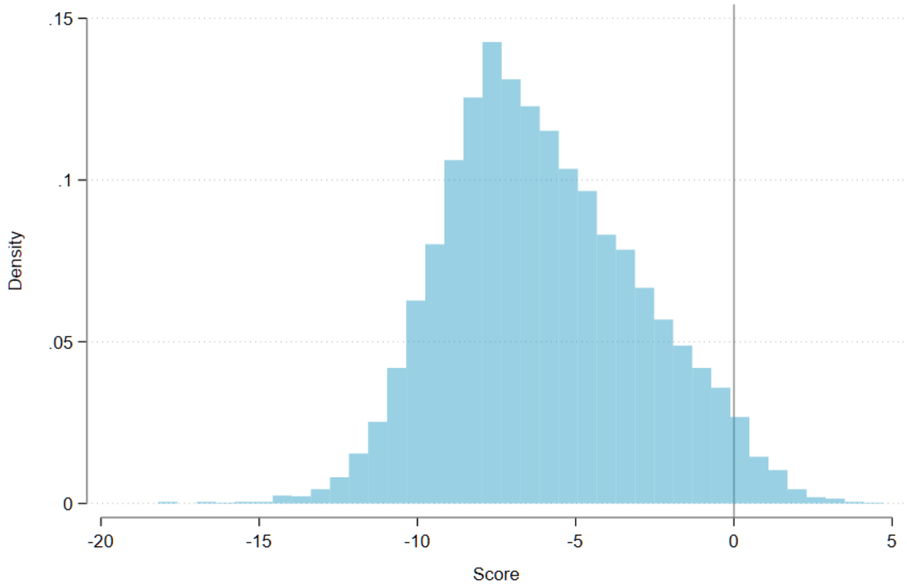


(h) Apgar score: post-intervention



Notes: Each dot represents the average outcome for a given deprivation score, centered at 0, and grouped into bins of 0.15 deprivation score units. The size of each marker is proportional to the number of observations in the bin. Observations within a bandwidth of $h = 2.8$ are used to fit the linear trends, which are estimated using individual-level data on either side of the threshold. Blue dots represent individuals in non-deprived neighborhoods, and red dots represent those in deprived neighborhoods. Panels (a), (c), (e), and (g) show outcomes from the pre-intervention period (2008), while panels (b), (d), (f), and (h) show outcomes from the post-intervention period (2009-2011). The vertical lines at $h = 1.4$ indicate the optimal bandwidth (see Section 4).

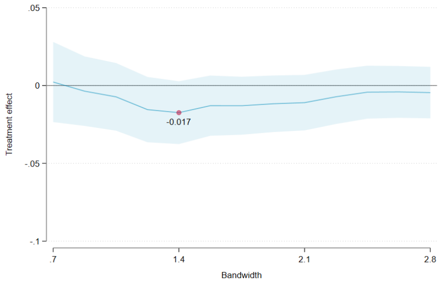
Figure A.5: Histogram of Deprivation Scores



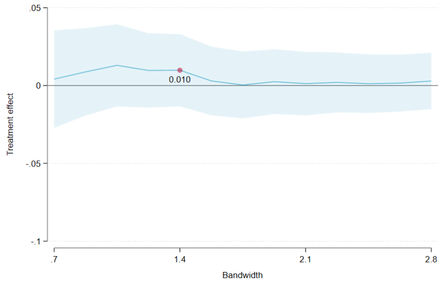
Notes: This figure shows the distribution of deprivation scores across neighborhoods, with each neighborhood represented as a single data point. Deprivation scores are time-invariant for the 20082011 period and have been normalized such that the cut-off value for categorizing neighborhoods as deprived is set to zero.

Figure A.6: Sensitivity of Treatment Effects to Bandwidth Selection

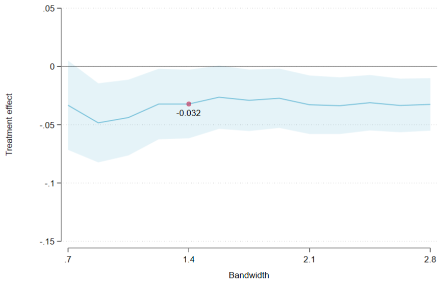
(a) Boys: low birthweight



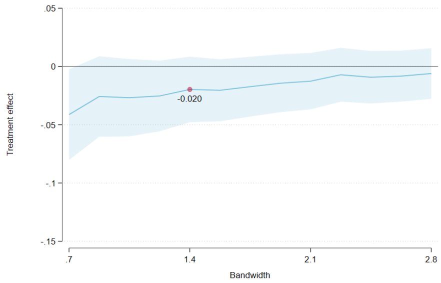
(b) Girls: low birthweight



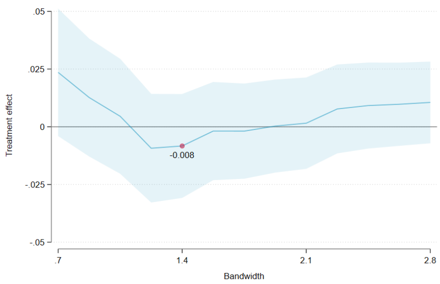
(c) Boys: small for gestational age



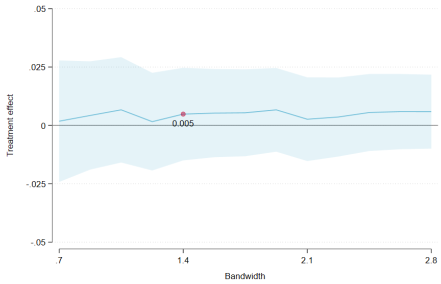
(d) Girls: small for gestational age



(e) Boys: preterm



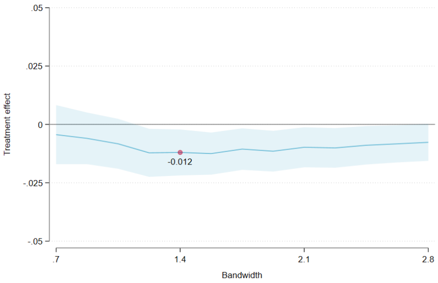
(f) Girls: preterm



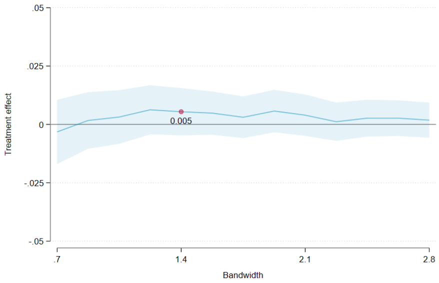
Notes: Panels (g) to (j) displayed on the next page

Figure A.6: Sensitivity of Treatment Effects to Bandwidth Selection

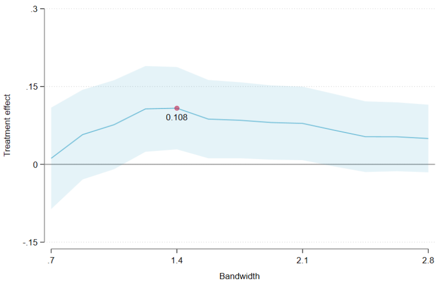
(g) Boys: Apgar <7



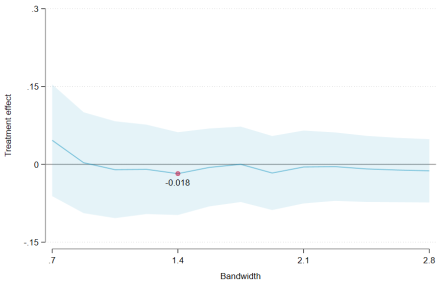
(h) Girls: Apgar <7



(i) Boys: summary index



(j) Girls: summary index



NNotes: This figure shows the sensitivity of treatment effect estimates to different bandwidth choices. Blue lines depict the treatment effect for each outcome as the bandwidth varies, while the shaded areas represent the 95% confidence intervals. The red dot indicates the estimate from the preferred specification, using a bandwidth of $h = 1.4$.

Table A.1: Annual Number of Births by Neighborhood

	(1) N	(2) Mean	(3) SD	(4) P25	(5) Median	(6) P75
All births	6,773	93.28	118	23	57	117
Main sample	538	175.51	200.8	54	108	206

Notes: This table summarizes the annual number of live births by neighborhood during the period 20082011, based on data from Perined. The "All births" row includes all neighborhoods with at least one live birth during the study period, while the "Main sample" row only includes neighborhoods that meet the criteria for inclusion in the main analysis. Summary statistics are provided for the number of births across neighborhoods, including the mean, standard deviation (SD), 25th percentile (P25), median, and 75th percentile (P75).

Table A.2: Sample Size by Treatment Arm and Period

	Boys		Girls	
	(1) Deprived	(2) Non-deprived	(3) Deprived	(4) Non-deprived
Pre-intervention (2008)	7,401	4,242	7,249	3,941
Post-intervention (2009-2011)	23,488	13,276	22,297	12,531
Total	48,407		46,018	

Notes: This table reports the sample size by treatment status (deprived vs. non-deprived neighborhoods), sex (boys vs. girls), and period (pre-intervention vs. post-intervention). The pre-intervention period covers births in 2008, while the post-intervention period covers births between 2009 and 2011. The total sample size is reported at the bottom for each group.

Table A.3: Optimal Bandwidth

	(1) Low birthweight	(2) Small for gestational age	(3) Preterm	(4) Apgar <7
Boys	1.51	1.50	1.57	1.81
Girls	1.13	1.33	1.16	1.12

Notes: This table reports the mean square error (MSE) optimal bandwidths calculated following the approach proposed by Calonico et al. (2020). The final bandwidth used in our analysis, $h^* = 1.4$, is the average of the bandwidths presented in this table. These bandwidths are used to estimate the treatment effects.

Table A.4: Effect of Deprivation Payment on Continuous Measures of Birth Outcomes

	(1)	(2)	(3)	(4)
	Birthweight	Birthweight centile	Gestational age	Apgar score
Treatment effect	49.63**	3.160**	0.155	0.0760**
	(23.90)	(1.252)	(0.581)	(0.0369)
Mean DV	3,406.942	45.789	275.435	9.579
N	48,407	48,407	48,407	48,407

(a) Boys

	(1)	(2)	(3)	(4)
	Birthweight	Birthweight centile	Gestational age	Apgar score
Treatment effect	12.10	0.644	0.0752	-0.0467
	(23.66)	(1.095)	(0.617)	(0.0347)
Mean DV	3,283.896	45.449	275.980	9.653
N	46,018	46,018	46,018	46,018

(b) Girls

Notes: This table presents estimates of the effect of deprivation payments on four continuous measures of health at birth. The dependent variables are: birth weight (measured in grams), birth weight centile (ranging from 1 to 100), gestational age (measured in days), and Apgar score (ranging from 1 to 10). The estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * $p< 0.1$, ** $p< 0.05$, *** $p< 0.01$.

Table A.5: Effect of Deprivation Payment on Parental Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Primiparous	Birth order	Maternal age	Dutch mother	Paternal age	Dutch father	Standardized income
Treatment effect	0.00398	-0.00997	0.267	0.0184	0.0838	0.0149	-0.903
	(0.0213)	(0.0426)	(0.193)	(0.0194)	(0.309)	(0.0238)	(409.0)
Mean DV	0.478	2.385	29.581	0.594	33.796	0.570	20393.016
N	48,407	48,407	48,407	48,407	43,683	43,682	46,828

(a) Boys

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Primiparous	Birth order	Maternal age	Dutch mother	Paternal age	Dutch father	Standardized income
Treatment effect	0.0210	0.0163	0.199	-0.0123	0.130	-0.0205	1017.3*
	(0.0194)	(0.0404)	(0.204)	(0.0192)	(0.288)	(0.0219)	(538.3)
Mean DV	0.475	2.404	29.519	0.596	33.769	0.572	20315.973
N	46,016	46,018	46,018	46,017	41,419	41,419	44,458

(b) Girls

Notes: This table presents estimates of the effect of deprivation payments on key covariates to validate the studys empirical approach. Panel (a) reports estimates for boys, and panel (b) for girls. Each column represents a separate estimate from equation 1. The covariates include primiparous status (dummy), birth order, maternal age, maternal Dutch nationality (dummy), paternal age, paternal Dutch nationality (dummy), and standardized disposable household income in the year before delivery. Disposable income consists of total household income, including income transfers, net of income taxes, and contributionssuch as social security contributions and health insurance premiums. The OECD modified equivalence scale (Organisation for Economic Co-operation and Development, 2008) is used to standardize for the household size in which the mother is registered. The estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * $p< 0.1$, ** $p< 0.05$, *** $p< 0.01$.

Table A.6: Effect of Deprivation Payment on Birth Outcomes in Placebo Years

	(1)	(2)	(3)	(4)	(5)
	Low birthweight	Small for gestational age	Preterm	Apgar <7	Summary index
Treatment x 2007	-0.00397	-0.0353*	-0.00145	-0.00310	0.0552
	(0.0150)	(0.0212)	(0.0167)	(0.00765)	(0.0570)
Mean DV	0.071	0.160	0.086	0.015	0.000
N	19,773	19,773	19,773	19,773	19,773

(a) Boys

	(1)	(2)	(3)	(4)	(5)
	Low birthweight	Small for gestational age	Preterm	Apgar <7	Summary index
Treatment x 2007	-0.0187	-0.0174	-0.0117	-0.0106	0.0964
	(0.0181)	(0.0233)	(0.0168)	(0.00838)	(0.0701)
Mean DV	0.078	0.161	0.072	0.015	0.000
N	18,834	18,834	18,834	18,834	18,834

(b) Girls

Notes: This table presents estimates of the effect of deprivation payments on key birth outcomes using data from the pre-intervention period (20062007). The dependent variables in columns (1)-(4) are binary indicators for: low birth weight (birth weight <2,500g), small for gestational age (birth weight below the 10th percentile, adjusted for gestational age and sex), preterm birth (gestational age <37 weeks), and low Apgar score (<7 at 5 minutes). Column (5) reports estimates for a summary index aggregating these outcomes into a single measure, standardized to have a mean of 0 and a standard deviation of 1. The estimates are based on equation 1, with standard errors clustered at the neighborhood level to account for within-neighborhood correlations in outcomes. Statistical significance is denoted by * p< 0.1, ** p< 0.05, *** p< 0.01.

Table A.7: Effect of Deprivation Payment on Birth Outcomes, Removing Observations +/- 1 Week from January 1, 2009

	(1)	(2)	(3)	(4)	(5)
	Low birthweight	Small for gestational age	Preterm	Apgar <7	Summary index
Treatment effect	-0.0162	-0.0316**	-0.00685	-0.0125**	0.105***
	(0.0103)	(0.0150)	(0.0115)	(0.00514)	(0.0407)
Mean DV	0.065	0.148	0.081	0.018	0.000
N	48,014	48,014	48,014	48,014	48,014

(a) Boys

	(1)	(2)	(3)	(4)	(5)
	Low birthweight	Small for gestational age	Preterm	Apgar <7	Summary index
Treatment effect	0.0101	-0.0184	0.00607	0.00575	-0.0224
	(0.0120)	(0.0143)	(0.0102)	(0.00525)	(0.0411)
Mean DV	0.077	0.152	0.073	0.014	0.000
N	45,668	45,668	45,668	45,668	45,668

(b) Girls

Notes: This table presents estimates of the effect of deprivation payments on key birth outcomes, excluding observations for births that occurred between 24 December 2008 and 7 January 2009. The dependent variables in columns (1)-(4) are binary indicators for: low birth weight (birth weight <2,500g), small for gestational age (birth weight below the 10th percentile, adjusted for gestational age and sex), preterm birth (gestational age <37 weeks), and low Apgar score (<7 at 5 minutes). Column (5) reports estimates for a summary index aggregating these outcomes into a single measure, standardized to have a mean of 0 and a standard deviation of 1. The estimates are based on equation 1, with standard errors clustered at the neighborhood level to account for within-neighborhood correlations in outcomes. Statistical significance is denoted by * p< 0.1, ** p< 0.05, *** p< 0.01.

Table A.8: Effect of Deprivation Payment on Birth Outcomes, Removing Observations +/- 2 Week from January 1, 2009

	(1)	(2)	(3)	(4)	(5)
	Low birthweight	Small for gestational age	Preterm	Apgar <7	Summary index
Treatment effect	-0.0151	-0.0327**	-0.00525	-0.0125**	0.102**
	(0.0105)	(0.0155)	(0.0117)	(0.00522)	(0.0419)
Mean DV	0.065	0.148	0.081	0.019	-0.000
N	47,563	47,563	47,563	47,563	47,563

(a) Boys

	(1)	(2)	(3)	(4)	(5)
	Low birthweight	Small for gestational age	Preterm	Apgar <7	Summary index
Treatment effect	0.0114	-0.0178	0.00868	0.00599	-0.0295
	(0.0122)	(0.0146)	(0.0103)	(0.00529)	(0.0420)
Mean DV	0.077	0.153	0.073	0.014	-0.000
N	45,242	45,242	45,242	45,242	45,242

(b) Girls

Notes: This table presents estimates of the effect of deprivation payments on key birth outcomes, excluding observations for births that occurred between 17 December 2008 and 14 January 2009. The dependent variables in columns (1)(4) are binary indicators for: low birth weight (birth weight <2,500g), small for gestational age (birth weight below the 10th percentile, adjusted for gestational age and sex), preterm birth (gestational age <37 weeks), and low Apgar score (<7 at 5 minutes). Column (5) reports estimates for a summary index aggregating these outcomes into a single measure, standardized to have a mean of 0 and a standard deviation of 1. Following Kling et al. (2007), the summary index is constructed as the sum of z-scores of its components, with each component oriented so that higher scores indicate more beneficial outcomes. A positive treatment effect on the summary index thus reflects improved health at birth. The estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * p< 0.1, ** p< 0.05, *** p< 0.01.

Table A.9: Effect of Deprivation Payment on Placebo Outcomes

	(1)	(2)
	Female	Multiple birth
Treatment effect	0.00642	-0.00160
	(0.0139)	(0.00311)
Mean DV	0.487	0.016
N	94,425	94,425

Notes: This table presents estimates of the effect of deprivation payments on two placebo outcomes: female (column 1) and multiple births (column 2). Both outcomes are binary indicator. The estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * p< 0.1, ** p< 0.05, *** p< 0.01.

Table A.10: Sensitivity of Treatment Effects to Alternative Specification of the Summary Index of Health at Birth

	Boys			Girls		
	(1)	(2)	(3)	(4)	(5)	(6)
	Z-scores, discrete	Z-scores, continuous	PCA	Z-scores, discrete	Z-scores, continuous	PCA
Treatment effect	0.108***	0.100**	0.0899**	-0.0179	-0.00233	-0.0166
	(0.0405)	(0.0395)	(0.0414)	(0.0407)	(0.0425)	(0.0410)
Mean DV	-0.000	-0.000	0.000	0.000	-0.000	0.000
N	48,407	48,407	48,407	46,018	46,018	46,018

Notes: PCA = principal component analysis. This table presents estimates of the effect of deprivation payments on different specifications of the summary index for health at birth. Columns (1) and (4) replicate the treatment effects on the binary summary index reported in Table 2. The summary index in columns (2) and (5) is constructed analogously, but using continuous measures of birth outcomes instead of binary indicators. Columns (3) and (6) report an alternative specification of the summary index, constructed using the first component from a principal component analysis (PCA) of the four main birth outcome measures, following the approach described in Cygan-Rehm and Karbownik (2022). Following Kling et al. (2007), all summary indices are constructed as the sum of z-scores of their components, with each component oriented so that higher scores indicate more beneficial outcomes. A positive treatment effect on the summary index thus reflects improved health at birth. The estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Sensitivity of Treatment Effects to Alternative Estimation Models: Boys

	(1)	(2)	(3)	(4)	(5)	(6)
			Low birthweight			
Treatment effect	-0.0174*	-0.0181*	-0.0178*	-0.0179*	-0.0148	-0.0152
	(0.0103)	(0.0102)	(0.0102)	(0.0102)	(0.00968)	(0.00974)
Mean DV	0.065	0.065	0.065	0.065	0.065	0.065
N	48,407	46,827	46,827	46,827	46,827	46,827

	(1)	(2)	(3)	(4)	(5)	(6)
			Small for gestational age			
Treatment effect	-0.0322**	-0.0308**	-0.0308**	-0.0311**	-0.0291*	-0.0295*
	(0.0150)	(0.0155)	(0.0154)	(0.0154)	(0.0151)	(0.0150)
Mean DV	0.148	0.148	0.148	0.148	0.148	0.148
N	48,407	46,827	46,827	46,827	46,827	46,827

	(1)	(2)	(3)	(4)	(5)	(6)
			Preterm			
Treatment effect	-0.00831	-0.00755	-0.00751	-0.00743	-0.00412	-0.00440
	(0.0115)	(0.0116)	(0.0117)	(0.0117)	(0.0109)	(0.0110)
Mean DV	0.081	0.081	0.081	0.081	0.081	0.081
N	48,407	46,827	46,827	46,827	46,827	46,827

	(1)	(2)	(3)	(4)	(5)	(6)
			Apgar <7			
Treatment effect	-0.0120**	-0.0130**	-0.0132***	-0.0131**	-0.0128**	-0.0129**
	(0.00502)	(0.00507)	(0.00506)	(0.00510)	(0.00513)	(0.00510)
Mean DV	0.019	0.019	0.019	0.019	0.019	0.019
N	48,407	46,827	46,827	46,827	46,827	46,827

	(1)	(2)	(3)	(4)	(5)	(6)
			Summary index			
Treatment effect	0.108***	0.110***	0.110***	0.110***	0.0973**	0.0990**
	(0.0405)	(0.0409)	(0.0410)	(0.0412)	(0.0391)	(0.0390)
Mean DV	-0.000	0.000	0.000	0.000	0.000	0.000
N	48,407	46,827	46,827	46,827	46,827	46,827
No missing info	No	Yes	Yes	Yes	Yes	Yes
Month and year FE	No	No	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	Yes	Yes	Yes
Maternal health controls	No	No	No	No	Yes	Yes
Demographic controls	No	No	No	No	No	Yes

Notes: This table presents estimates of the effect of deprivation payments on the main health outcomes, tested across different samples, controls, and fixed effects specifications. The dependent variables in the four top panels are binary indicators for: low birth weight (birth weight <2,500g), small for gestational age (birth weight below the 10th percentile, adjusted for gestational age and sex), preterm birth (gestational age <37 weeks), and low Apgar score (<7 at 5 minutes). The fifth one reports estimates for a summary index aggregating these outcomes into a single measure, standardized to have a mean of 0 and a standard deviation of 1. Following Kling et al. (2007), the summary index is constructed as the sum of z-scores of its components, with each component oriented so that higher scores indicate more beneficial outcomes. A positive treatment effect on the summary index thus reflects improved health at birth. Column (1) replicates the main specification from Table 2. Columns (2) through (6) progressively incorporate additional robustness checks. In columns (2) to (6), observations with missing information for any control variables are excluded. Column (3) adds year and month fixed effects. Column (4) includes municipality fixed effects. Column (5) controls for health risk factors, including parity, birth order, and maternal age at birth. Finally, column (6) includes additional socioeconomic and demographic covariates: maternal Dutch nationality (dummy), father's age at birth, father's Dutch nationality (dummy), and disposable household income (in euros). The estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * $p< 0.1$, ** $p< 0.05$, *** $p< 0.01$.

Table A.12: Sensitivity of Treatment Effects to Alternative Estimation Models: Girls

	(1)	(2)	(3)	(4)	(5)	(6)
			Low birthweight			
Treatment effect	0.00984	0.00800	0.00808	0.00721	0.00536	0.00675
	(0.0118)	(0.0120)	(0.0120)	(0.0120)	(0.0113)	(0.0113)
Mean DV	0.077	0.077	0.077	0.077	0.077	0.077
N	46,018	44,455	44,455	44,455	44,455	44,455

	(1)	(2)	(3)	(4)	(5)	(6)
			Small for gestational age			
Treatment effect	-0.0197	-0.0217	-0.0217	-0.0228	-0.0241*	-0.0217
	(0.0143)	(0.0145)	(0.0145)	(0.0147)	(0.0144)	(0.0144)
Mean DV	0.152	0.151	0.151	0.151	0.151	0.151
N	46,018	44,455	44,455	44,455	44,455	44,455

	(1)	(2)	(3)	(4)	(5)	(6)
			Preterm			
Treatment effect	0.00485	0.00288	0.00293	0.00247	0.000700	0.00147
	(0.0101)	(0.0106)	(0.0106)	(0.0106)	(0.0103)	(0.0103)
Mean DV	0.073	0.072	0.072	0.072	0.072	0.072
N	46,018	44,455	44,455	44,455	44,455	44,455

	(1)	(2)	(3)	(4)	(5)	(6)
			Apgar <7			
Treatment effect	0.00544	0.00566	0.00562	0.00518	0.00490	0.00503
	(0.00516)	(0.00522)	(0.00522)	(0.00522)	(0.00522)	(0.00522)
Mean DV	0.014	0.014	0.014	0.014	0.014	0.014
N	46,018	44,455	44,455	44,455	44,455	44,455

	(1)	(2)	(3)	(4)	(5)	(6)
			Summary index			
Treatment effect	-0.0179	-0.0109	-0.0109	-0.00645	0.00116	-0.00493
	(0.0407)	(0.0409)	(0.0410)	(0.0409)	(0.0390)	(0.0388)
Mean DV	0.000	0.000	0.000	0.000	0.000	0.000
N	46,018	44,455	44,455	44,455	44,455	44,455
No missing info	No	Yes	Yes	Yes	Yes	Yes
Month and year FE	No	No	Yes	Yes	Yes	Yes
Municipality FE	No	No	No	Yes	Yes	Yes
Maternal health controls	No	No	No	No	Yes	Yes
Demographic controls	No	No	No	No	No	Yes

Notes: This table presents estimates of the effect of deprivation payments on the main health outcomes, tested across different samples, controls, and fixed effects specifications. The dependent variables in the four top panels are binary indicators for: low birth weight (birth weight <2,500g), small for gestational age (birth weight below the 10th percentile, adjusted for gestational age and sex), preterm birth (gestational age <37 weeks), and low Apgar score (<7 at 5 minutes). The fifth one reports estimates for a summary index aggregating these outcomes into a single measure, standardized to have a mean of 0 and a standard deviation of 1. Following Kling et al. (2007), the summary index is constructed as the sum of z-scores of its components, with each component oriented so that higher scores indicate more beneficial outcomes. A positive treatment effect on the summary index thus reflects improved health at birth. Column (1) replicates the main specification from Table 2. Columns (2) through (6) progressively incorporate additional robustness checks. In columns (2) to (6), observations with missing information for any control variables are excluded. Column (3) adds year and month fixed effects. Column (4) includes municipality fixed effects. Column (5) controls for health risk factors, including parity, birth order, and maternal age at birth. Finally, column (6) includes additional socioeconomic and demographic covariates: maternal Dutch nationality (dummy), father's age at birth, father's Dutch nationality (dummy), and disposable household income (in euros). The estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * $p< 0.1$, ** $p< 0.05$, *** $p< 0.01$.

Table A.13: Effect of Deprivation Payment on Birth Outcomes Among Stayers

	(1)	(2)	(3)	(4)	(5)
	Low birthweight	SGA	Preterm	Apgar < 7	Summary index
Treatment effect	-0.0179*	-0.0286*	-0.0118	-0.0114**	0.112***
	(0.0106)	(0.0157)	(0.0117)	(0.00493)	(0.0413)
Mean DV	0.063	0.147	0.079	0.018	0.000
N	43,963	43,963	43,963	43,963	43,963

(a) Boys

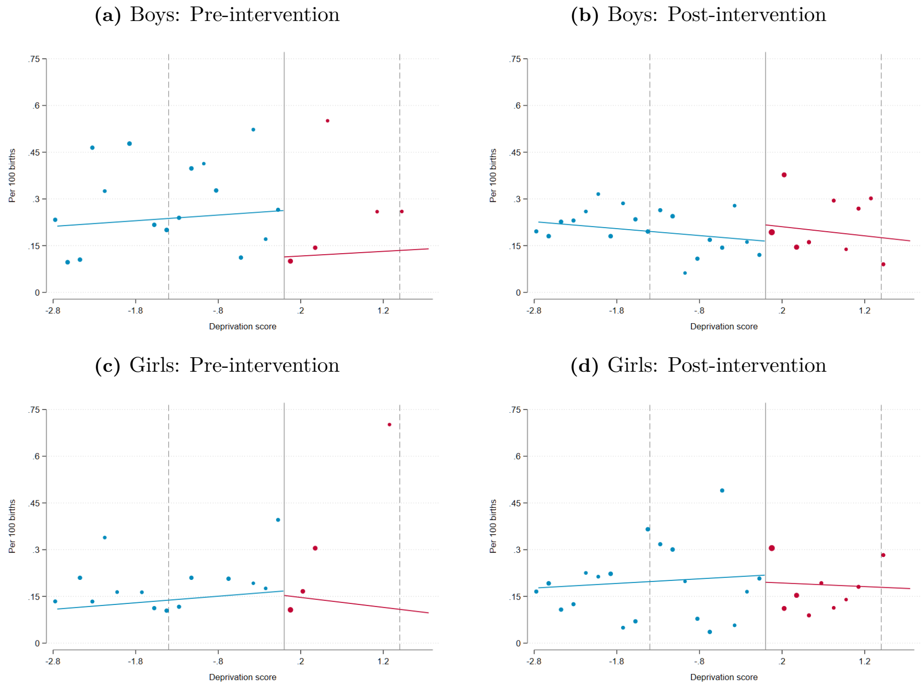
	(1)	(2)	(3)	(4)	(5)
	Low birthweight	SGA	Preterm	Apgar < 7	Summary index
Treatment effect	0.00519	-0.0278*	0.00545	0.00462	-0.00117
	(0.0123)	(0.0147)	(0.0105)	(0.00483)	(0.0417)
Mean DV	0.075	0.149	0.071	0.014	-0.000
N	41,809	41,809	41,809	41,809	41,809

(b) Girls

Notes: This table presents estimates of the effect of deprivation payments on key birth outcomes, restricting the sample to infants whose mothers did not relocate in the 18 months before giving birth. The dependent variables in columns (1)-(4) are binary indicators for: low birth weight (birth weight <2,500g), small for gestational age (birth weight below the 10th percentile, adjusted for gestational age and sex), preterm birth (gestational age <37 weeks), and low Apgar score (<7 at 5 minutes). Column (5) reports estimates for a summary index aggregating these outcomes into a single measure, standardized to have a mean of 0 and a standard deviation of 1. TThe estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * p< 0.1, ** p< 0.05, *** p< 0.01.

Appendix B: Early Neonatal Mortality

Figure B.1: Early Neonatal Mortality by Deprivation Score Bins



(a) Boys: Pre-intervention

(b) Boys: Post-intervention

(c) Girls: Pre-intervention

(d) Girls: Post-intervention

Notes: Each dot represents the average early neonatal mortality rate for a given deprivation score, centered at 0, and grouped into bins of 0.15 deprivation score units. To comply with the guidelines by Statistics Netherlands, we do not plot dots with an average of zero. The size of each marker is proportional to the number of observations in the bin. Observations within a bandwidth of $h = 2.8$ are used to fit the linear trends, which are estimated using individual-level data on either side of the threshold. Blue dots represent individuals in non-deprived neighborhoods, and red dots represent those in deprived neighborhoods. Panels (a) and (c) show outcomes from the pre-intervention period (2008), while panels (b) and (d) show outcomes from the post-intervention period (2009-2011). The vertical lines at $h = 1.4$ indicate the optimal bandwidth.

Table B.1: Effect of Deprivation Payment on Early Neonatal Mortality

	Boys	Girls
	(1)	(2)
	Early neonatal mortality	
Treatment effect	0.00165	0.00205
	(0.00153)	(0.00171)
Mean DV	0.0019	0.0019
N	48,407	46,018

Notes: This table presents estimates of the effect of deprivation payments on early neonatal mortality, defined as deaths occurring between days 1 and 7 after birth. Each column represents a separate estimate from equation 1. The estimates are based on equation 1, with standard errors clustered at the neighborhood level, which is the unit of treatment. Statistical significance is denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.