UNIVERSIDADE FEDERAL DO PARANÁ

LEONARDO MATSUNO DA FROTA

ESSAYS ON NEWBORN´S HEALTH AND SOCIECONOMIC CONDITIONS

CURITIBA

2022

LEONARDO MATSUNO DA FROTA

ESSAYS ON NEWBORN´S HEALTH AND SOCIECONOMIC CONDITIONS

Tese para a obtenção do título de Doutor no Programa de Pós Graduação em Desenvolvimento Econômico do Setor de Ciências Sociais Aplicadas da Universidade Federal do Paraná
Orientador: Prof Dr. Marcos Minoru Hasegawa
Coorientador: Prof Dr. Paulo de Andrade Jacinto

CURITIBA

2022

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIAS SOCIAIS APLICADAS

Frota, Leonardo Matsuno da
Essays on newborn´s health and socieconomic conditions / Leonardo Matsuno da Frota. – Curitiba, 2022.
1 recurso on-line : PDF.

Tese (Doutorado) – Universidade Federal do Paraná, Setor de Ciências Sociais Aplicadas, Programa de Pós-Graduação em Desenvolvimento Econômico.
Orientador: Prof. Dr. Marcos Minoru Hasegawa.
Coorientador: Prof. Dr. Paulo de Andrade Jacinto.

1. Desenvolvimento econômico. 2. Cuidado pré-natal. 3. Economia da saúde. 4. Mortalidade infantil. 5. Recém-nascido – Saúde e higiene. I. Hasegawa, Marcos Minoru. II. Jacinto, Paulo de Andrade. III. Universidade Federal do Paraná. Programa de Pós-Graduação em Desenvolvimento Econômico. IV. Título.

Bibliotecária: Maria Lidiane Herculano Graciosa CRB-9/2008

MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS SOCIAIS E APLICADAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO DESENVOLVIMENTO ECONÔMICO - 40001016024P0

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação DESENVOLVIMENTO ECONÔMICO da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **LEONARDO MATSUNO DA FROTA** intitulada: **ESSAYS ON NEWBORN´S HEALTH AND SOCIECONOMIC CONDITIONS**, sob orientação do Prof. Dr. MARCOS MINORU HASEGAWA, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 23 de Fevereiro de 2022.

Assinatura Eletrônica
04/03/2022 10:48:12.0
MARCOS MINORU HASEGAWA
Presidente da Banca Examinadora

Assinatura Eletrônica
09/03/2022 11:43:03.0
LIVIA MADEIRA TRIACA
Avaliador Externo (UNIVERSIDADE FEDERAL DO RIO GRANDE)

Assinatura Eletrônica
03/03/2022 12:14:07.0
VICTOR RODRIGUES DE OLIVEIRA
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica
05/03/2022 23:13:08.0
ANDERSON MOREIRA ARISTIDES DOS SANTOS
Avaliador Externo (UNIVERSIDADE FEDERAL DE ALAGOAS)

Assinatura Eletrônica
03/03/2022 11:18:43.0
PAULO DE ANDRADE JACINTO
Coorientador(a) (UNIVERSIDADE FEDERAL DO PARANÁ)

Este trabalho é dedicado a todos os recém-nascidos. Que suas vidas sejam livres de sofrimento e das causas do sofrimento.

AGRADECIMENTOS

A minha mãe Maria Helena e meu pai Ruy por me darem todas as condições materiais e afetivas desde o momento que nasci até o presente momento. Ao meu irmão André por ser esse eterno figura.

Aos meus professores orientadores Prof. Dr. Marcos Hasegawa e Prof. Dr. Paulo Jacinto pelo anos de paciência e dedicação a esse trabalho. Também agradeço antecipadamente aos membros da banca Prof. Dr. Anderson Santos, Profa. Dra. Livia Madeira Triaca e Prof. Dr. Victor Rodrigues de Oliveira pela disponibilidade de avaliar esse trabalho.

Aos meus colegas de pós-graduação Edson Medeiros, Júlio Cateia, Caio Matteucci, Lucas Lopes, Auberth Venson, Victor Cruz e Silva, Lucas Casonatto, Luís Gustavo, Vitor Pestana, Eduardo Thome, João Tonin, Diego Liberato Souza, Luana Falcão, Ihorana Cuco, Carol Novaes, Helton Canguçu e também todos os que não foram citados nominalmente. Sem amigos, um doutorado é muito mais chato e difícil.

Especialmente, agradeço a Amanda Gentil, minha parceira de muitas jornadas. Sem o seu apoio, companheirismo e amor seria impossível ter terminado essa trabalho.

Aos lamas do Chagdud Gonpa, fontes incessantes de sabedoria nesses tempos estranhos. Aos meus alunos de graduação que são uma contínua motivação para continuar estudando. Por fim, agradeço a todos aqueles que não foram citados, mas de que alguma forma contribuíram para que esse trabalho fosse possível.

"Even if you must die tomorrow, study today,
Though you may not become a sage in this life,
Your knowledge will be preserved for the future
Just like wealth deposited and then reclaimed.
(Sakya Pandita, The Precious Treasury of Elegant Sayings)

RESUMO

Esta tese é composta por três ensaios sobre a relação entre a saúde do recém-nascido e as condições socioeconômicas. O primeiro ensaio diz respeito à identificação dos fatores de risco associados à mortalidade infantil. Usando 2,9 milhões de observações dos dados do Sistema Único de Saúde (SUS) de 2017, estimamos um conjunto de diferentes modelos de aprendizado de máquina para prever quais bebês têm maior risco de não sobreviver ao primeiro ano de vida. Descobrimos que, pela medida do índice de concordância, os modelos Survival Support Vector Machines, Extreme Gradient Boosting e Random Survival Forest podem gerar previsões de mortalidade muito precisas. Além disso, os modelos de aprendizado de máquina indicam que fatores como cesarianas, semanas gestacionais e baixo peso afetam a mortalidade de forma não linear. O segundo ensaio analisou o impacto do pré-natal no peso do recém-nascido, utilizando dados de uma amostra de 7,3 milhões de nascimentos entre 2015 e 2017 do Sistema Único de Saúde (SUS). A estrutura de correspondência do escore de propensão foi utilizada para avaliar o impacto do pré-natal inadequado na probabilidade de baixo peso ao nascer (<2500g). Além disso, um modelo de variável instrumental com efeitos fixos foi utilizada para mensurar o impacto da assistência pré-natal no peso ao nascer. Nossos achados são que um número inadequado de consultas de pré-natal aumenta as chances de baixo peso (<2500g) e muito baixo peso (<1500g) para os recém-nascidos da amostra. Além disso, cada consulta de pré-natal tem efeito médio positivo no peso ao nascer e cada mês de atraso no pré-natal tem efeito médio negativo. O terceiro ensaio examinou o impacto da cesariana em recém-nascidos de gestações pélvicas utilizando uma amostra de 28 mil partos do Sistema Único de Saúde (SUS). Um método de ponderação de probabilidade inversa de tratamento foi usado para medir o impacto da cesariana nos escores de APGAR e mortalidade infantil no primeiro ano de vida, abordando o viés de autoseleção inerente a esse cenário. Nossos achados são que, para bebês pélvicos, ter uma cesariana diminui a probabilidade de ter baixos escores de APGAR e morte. Não há evidência de impacto na probabilidade de nascimento com baixo peso (<2500g).

Palavras-chaves: Pré-Natal; Mortalidade Infantil; Saúde do Recém-nascido; Economia da Saude; Cesárea;

ABSTRACT

This thesis consists of three essays on the relationship between the newborn´s health and socioeconomic conditions. The first essay concerns the identification of risk factors associated with infant mortality. Using 2.9 million observations from the Brazilian Unique Health System (SUS) 2017 data, we estimated a set of different machine learning models to predict which infants have the highest risk of not surviving the first year of life. We found that by the concordance index measure, the Survival Support Vector Machines, the Extreme Gradient Boosting , and the Random Survival Forest models can generate very accurate mortality predictions.Also, the machine learning models indicate that factors such as cesarean sections, gestational weeks, and low weight affect mortality nonlinearly. The second essay examined the impact of prenatal care on the newborn´s weight, using data from a sample of 7.3 million births between 2015 and 2017 from the Brazillian Unique Health System (SUS). A propensity score matching framework was used to assess the impact of inadequate prenatal care on the probability of low birth weight (<2500g). Also, a fixed-effects instrumental variable was used to measure the prenatal care impact on birth weight. Our findings are that an inadequate number of prenatal care visits increase the odds of low weight (<2500g) and very low birth weight (<1500g) for newborns in the sample. Also, each prenatal care visit has a positive mean effect in the birthweight and each delayed month in prenatal care has a negative mean effect. The third essay examined the impact of having a cesarean section (C-Section) on newborns born from breech pregnancies using a sample of 28 thousands births from the Brazillian Unique Health System (SUS). An inverse probability of treatment weighting method was used to measure the c-section impact on the infant´s APGAR scores and mortality in the first year of life, addressing the self selection bias inherent in this setting. Our findings are that, for breech babies, having a C-Section decreases the probability of having low APGAR scores and death in There is no evidence of impact in the probability of having low weight birth (<2500g).

Key-words: Prenatal Care; Child mortality; Newborn Health; Health Economics; Cesarean Section

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

SUMÁRIO

1 INFANT MORTALITY IN BRAZIL: A SURVIVAL ANALYSIS USING MACHINE LEARNING MODELS

1.1 ABSTRACT

The persistence of infant mortality in middle-income countries like Brazil is a critical health challenge of the 21st century. Health care policymakers increasingly use statistical methods such as survival analysis to identify factors associated with mortality rates. A common choice in survival analysis is the Cox proportional hazards model. It is argued that in the presence of non-proportional hazards, the Cox model has limitations. Machine learning models are efficient at prediction and are a methodological alternative to models with the proportional hazards assumption. Using 2.9 million observations from the Brazilian Unique Health System (SUS) 2017 data, we estimated a set of different machine learning models (Survival Support Vector Machines, Random Survival Forest, and Extreme Gradient Boosting) to predict which infants have the highest risk of not surviving the first year of life. We found that by the concordance index measure, the Survival Support Vector Machines (c-index: 0.84), the Extreme Gradient Boosting (c-index: 0.83), and the Random Survival Forest (c-index: 0.81) models can generate very accurate mortality predictions. However, the Cox model also achieves accurate mortality predictions (c-index: 0.83) despite the presence of non proportional hazards. The SHAP framework of interpretable machine learning was used to identify factors affecting Brazil's infant mortality rates. Factors such as cesarean sections and gestational weeks affect mortality nonlinearly and mean variable effects such as those found in standard regression models can be misleading. Finally, we argue that interpretable Machine Learning models can support policymakers in designing health frameworks that tackle the challenge of infant mortality in middle-income countries.

Keywords: Brazil; Newborns health; Infant mortality; Survival analysis; Machine learning; Random survival forest;

1.2 INTRODUCTION

The deaths of children are particularly tragic events, as they are early and, in most cases, preventable deaths. Indeed, infant mortality is an important indicator of a population´s health and well being. As such, one of the United Nations Sustainable Development Goals is to reduce global newborn and infant mortality rates (ASSEMBLY, 2015). Quality health care and better socioeconomic conditions are instrumental for achieving this objective. Furthermore, the emergence of data driven health care can be an important ally in supporting this goal by identifying risk factors and helping design efficient policy frameworks (GROSSGLAUSER; SANER, 2014).

In Brazil, a middle income country, the infant mortality rate shows a decreasing trend from 1980 (78.5 deaths per thousand live births) to 2015 (12.1 deaths per thousand live births), according to World Bank data. Nevertheless, these rates are still high relative to developed economies (less than 5 deaths per thousand) (BANK, 2021). An important reason as to why research on infant mortality continues to be relevant in development economics and health policy. The empirical literature on this field is vast but a particular method that is insightful in this theme is survival analysis.

Particularly in Brazil, there are several studies using survival analysis that shed light in different perspectives of infant mortality. For instance, a survival analysis in intensive care units identify low birth weight (below 2500g) as a major risk factors for neonatal deaths (RISSO; NASCIMENTO, 2010). Another studies found higher mortality among children with low birth weight (below 2500g), born in public hospitals, as well as mothers with less schooling , and with insufficient prenatal visits (CARDOSO et al., 2013) (PINHEIRO; PERES; D'ORSI, 2010) (GARCIA; FERNANDES; TRAEBERT, 2019). Studies also discuss health conditions such as sepsis and congenital heart diseases as major risk factor for newborns survival (LOPES et al., 2018) (FREITAS et al., 2019). The most compreehensive survival analysis study uses 17.6 million births between 2011 and 2018 identifying three newborn characteristics that drive infant mortality: premature births (less than 37 gestational weeks), low weight and small for gestational age (babies that are below the 10th percentile weight for the same gestational weeks) (PAIXAO et al., 2021b). All these survival analysis studies in Brazil use Cox regression models that rely on the proportional hazards hypothesis. An assumption that is not warranted in every context and should be tested (ROYSTON; PARMAR, 2014) (GRAMBSCH; THERNEAU, 1994).

Machine Learning (ML) algorithms are a modelling alternative that can deal with non proportional hazards. In Netherlands, there is evidence that interpretable machine learning can provide efficient predictions and identification of risk factors for cancer mortality (MONCADA-TORRES et al., 2021). There is also evidence that Random Survival Forests can improve survival predictions on pacients with heart failures and cardiovascular diseases in general (MIAO et al., 2015) (MIAO et al., 2018). A study in Uganda uses a Random Survival Forest modeling strategy to identify the determinants of infant mortality and shows how the proportional hazards assumption diminishes the model robustness (NASEJJE; MWAMBI, 2017).

There are studies that use machine learning to discuss infant mortality in Brazil, but they do not use survival analysis methods. A study uses several ML algorithms and proposes a governance framework to identify the determinants of infant mortality (RAMOS et al., 2017). Another study uses data available in the gestational period to argue that it is possible to identify infants with high risk of mortality before birth (VALTER et al., 2019).In a small sample of 15 thousand births in Sao Paulo, there is also evidence that interpretable machine learning can identify newborns at high risk of death using public health databases (BELUZO et al., 2020). A more recent study with 1.2 million births in Sao Paulo finds that the extreme gradient boosting (XGBOOST) model has the best predictive performance in identifying infant mortality (BATISTA et al., 2021).

On a methodological level, our contribution is to show the efficiency of machine learning survival models in predicting infants with high risk of death, but also the prediction efficiency of the Cox model even when there is non-proportional hazards. Also to show how interpretable machine learning algorithms can assess non-linearities in the determinants of infant mortality. We also contribute in using micro-data of 2.9 million births from the Unique Health System (SUS), this is the first paper to use survival analysis with machine learning to assess infant mortality in Brazilian micro-data. Our main findings are that Survival Support Vector Machines (SSVM), Random Survival Forests (RSF) and Extreme Gradient Boosting (XGBOOST) models achieve a strong predictive performance (concordance index > 0.8) in identifying infants that died in the first year of life. However, only the Survival Support Vector Machines and the Extreme Gradient Boosting models beat the Cox regression benchmark performance.Furthermore, the SHAP[1] algorithm of interpretable machine learning shows non-linearities in the relationship between individual features such as gestational weeks and c section that are not explicit in typical survival models. We also discuss public policy implications and caveats in the development of predictive frameworks that help predict and identify risk factors associated with infant mortality.

1.3 DATA

We use two different datasets from the Brazilian Unified Health System (SUS) in this research. The Live Birth Information System (SINASC) and the Mortality Information System (SIM). The Live Birth Information System (SINASC) function collects and processes demographic and epidemiological data on newborn characteristics and mother characteristics. It is structured around the Live Birth Declaration (DN). The system is universal in the Brazilian territory, and it is expected that the professionals working in the health services or in the registry offices fill in the Live Birth Declaration (DN). The Mortality Information System (SIM) is a system of national epidemiological surveillance whose objective is to capture data on the country's deaths to provide information on mortality for all instances of the Brazilian health system. It is structured around the declaration of death (DO).

Our sample is created by matching information from the two above datasets using 2017 data. Each birth has an associated number in the Live Birth Information System - the "DN" code. If the infant was born and died in the same year, he will have a DN and a DO code. Then, to get infant mortality, we match the Mortality Information System with the Live Birth Information System using the "DN" code and "DO" code. In this way, we have a dataset

[1] the SHAP algorithm uses the game theoretical concept of Shapley values to explain the predictions of machine learning models (LUNDBERG; LEE, 2017).

TABLE 1 – Original Variables Description

Variable Name	Definition	Type	Source
DN	Infant ID	Numerical	Livebirth Dataset (SINASC)
Birth Location	Place of Birth	Categorical	Livebirth Dataset (SINASC)
Mother Age	Mother´s age in years	Numerical	Livebirth Dataset (SINASC)
Marital Status	Marital Status	Categorical	Livebirth Dataset (SINASC)
Schooling	Mother´s Education in levels	Categorical	Livebirth Dataset (SINASC)
Live children	Number of living children	Numerical	Livebirth Dataset (SINASC)
Number of dead children	Number of dead children	Numerical	Livebirth Dataset (SINASC)
Gestational Weeks	Gestational Weeks	Numerical	Livebirth Dataset (SINASC)
Parity	Type of Pregnancy (Unique; Double; Triple)	Categorical	Livebirth Dataset (SINASC)
C Section	Type of delivery: Vaginal or C-Section	Categorical	Livebirth Dataset (SINASC)
Prenatal Visits	Number of pre-natal care visits	Numerical	Livebirth Dataset (SINASC)
Sex	Infant Sex	Categorical	Livebirth Dataset (SINASC)
APGAR1	1st minute APGAR	Numerical	Livebirth Dataset (SINASC)
APGAR5	Fifth minute APGAR	Numerical	Livebirth Dataset (SINASC)
Race	Race/Ethnicity	Categorical	Livebirth Dataset (SINASC)
Birth Weight	birth weight in grams	Numerical	Livebirth Dataset (SINASC)
Genetic Anomaly	Genetic Anomaly	Categorical	Livebirth Dataset (SINASC)
Previous Gestations	Number of Previous Gestations	Numerical	Livebirth Dataset (SINASC)
Vaginal births	Number of Vaginal Births	Numerical	Livebirth Dataset (SINASC)
Cesarean births	Number of c-sections	Numerical	Livebirth Dataset (SINASC)
Induced labor	Induced Labor	Categorical	Livebirth Dataset (SINASC)
Fetus Presentation	Fetus position before Labor	Categorical	Livebirth Dataset (SINASC)
C section before start	C Section began before labor	Categorical	Livebirth Dataset (SINASC)
Death	Death in the first year of life	Binary	Mortality Dataset (SIM)
Birth Date	Date of Infant´s Birth	Calendar	Livebirth Dataset (SINASC)
Date of Death	Date of Infant´s Death	Calendar	Mortality Dataset (SIM)

Source: Prepared by the authors using Unified Health System (SUS) data.
Notes: The table describes the variable name that we adopted in our estimations, the definition of each variable based on SUS data, the variable type and the corresponding data source from the SUS.

containing newborn and mother characteristics and information regarding infant mortality in the first year of life - in this particular case, information from 2017. Table 7 shows the set of variables that will be used in the statistical analysis and their respective definitions on the original datasets.

An essential characteristic of our dataset is that we can have specific date information for all births and the death dates of the infants that did not survive the first year. These two variables allow us to create time-to-event indicators necessary to the proper usage of survival analysis models - our modeling choice to assess infant mortality in Brazil.

Since a substantial part of our original data is categoric, we perform several variable transformations to prepare the original data (table 1) for the statistical estimations. We also remove null observations and observations with infinite numeric values on the dataset due to measurement error. The transformed variables used in the models are shown in table 2 which describes the set of numeric variables and their summary statistics for live and dead infants and in table 3 which describes the distribution of births and deaths by each dummy variable.

TABLE 2 – Summary Statistics

Variable Name	Alive	Dead
Gestational Weeks	38.56	31.92
	(2.03)	(6.19)
Mother Age	26.74	26.66
	(6.69)	(7.32)
Prenatal Visits	8.03	5.91
	(2.75)	(2.99)
Observations	2646432	20310

Source: Prepared by the authors using Unified Health System (SUS) data.
Note: The table describes the mean and standard deviation (in parentheses) for the numeric variables used in the models. Columns divide the sample between infants who survived and did not survive the first year of life.

We create dummies for Genetic Anomaly (1 for genetic anomaly zero otherwise), Birth Place (1 for Hospital Birth and zero otherwise), Border (1 for border municipality birth and zero otherwise), Capital (1 for birth in a state capital and zero otherwise), Birth type (1 for c-section and zero for vaginal birth), Low APGAR 5 [2] (1 for low APGAR5 score and zero otherwise), Low Birth Weight (1 for low birth weight and zero otherwise), Mother Marital Status (1 for married and zero otherwise), Premature (1 for premature birth and zero otherwise), Mother Race (1 for white mother and zero otherwise), Schooling (1 if the mother went to college and zero if not), Baby Gender (1 for male infant and zero for female).

After the data transformations, our sample has 2.64 million births and 20310 deaths. The proportion of dead infants is less than 1% of the total births. There are differences in the proportions between live and dead infants that shed light on the risk factors for infant mortality. For instance, 14.91% of newborns with genetic anomalies do not survive, 28.5% of low APGAR5, and 6.39% of low weight babies also die during the first year of life. There are also differences in the means of numeric variables that indicate risk factors. The mean of prenatal visits is 8.03 for newborns who survived the first year and 5.91 for newborns who did not. Finally, the mean of gestational weeks is seven weeks lower for dead infants.

[2] the APGAR score is a test performed on newborns shortly after birth that assesses their general state and vitality. This assessment is done in the first minute of birth and is repeated again 5 minutes after delivery, taking into account baby characteristics such as heartbeat, color, breathing and natural reflexes (CASEY; MCINTIRE; LEVENO, 2001)

TABLE 3 – The distribution of births and deaths by features

Variable Name	Number of observations (proportion)		Total
	Alive N (%)	Dead N (%)	
Genetic Anomaly			
Genetic Anomaly	19828 (85.09)	3474 (14.91)	23302
Non-Genetic Anomaly	2626604 (99.36)	16836 (0.64)	2643440
Birth Place			
Hospital Birth	2623928 (99.24)	20090 (0.76)	2644018
Non-Hospital Birth	22504 (99.03)	220 (0.97)	22724
Border			
Border Municipality Birth	159978 (99.12)	1420 (0.88)	161398
Otherwise	2486454 (99.25)	18890 (0.75)	2505344
Capital			
State Capital Municipality Birth	636116 (99.24)	4890 (0.76)	641006
Otherwise	2010316 (99.24)	15420 (0.76)	2025736
Birth Type			
C-Section	1504026 (99.29)	10737 (0.71)	1514763
Vaginal Birth	1142406 (99.17)	9573 (0.83)	1151979
APGAR5			
Low APGAR5	19070 (71.50)	7600 (28.50)	26670
Normal APGAR5	2627362 (99.53)	12710 (0.48)	2640072
Birth Weight			
Low Weight	205584 (93.61)	14024 (6.39)	219608
Normal Weight	2440848 (99.74)	6286 (0.26)	2447134
Mother Marital Status			
Married	896252 (99.36)	5781 (0.64)	902033
Non-Married	1750180 (99.18)	14529 (0.82)	1764709
Premature			
Non-premature Birth	2524928 (99.69)	7766 (0.31)	2532694
Premature Birth	121504 (90.64)	12544 (9.36)	134048
Mother Race			
White Mother	1677205 (99.20)	13594 (0.80)	1690799
Otherwise	969227 (99.31)	6716 (0.69)	975943
Schooling			
Mother went to College	549756 (99.39)	3368 (0.61)	553124
Did not go to College	2096676 (99.20)	16942 (0.80)	2113618
Baby Gender			
Female	1355840 (99.19)	11117 (0.81)	1366957
Male	1290592 (99.29)	9193 (0.71)	1299785
Observations	2646432 (99.23)	20310 (0.72)	2666742

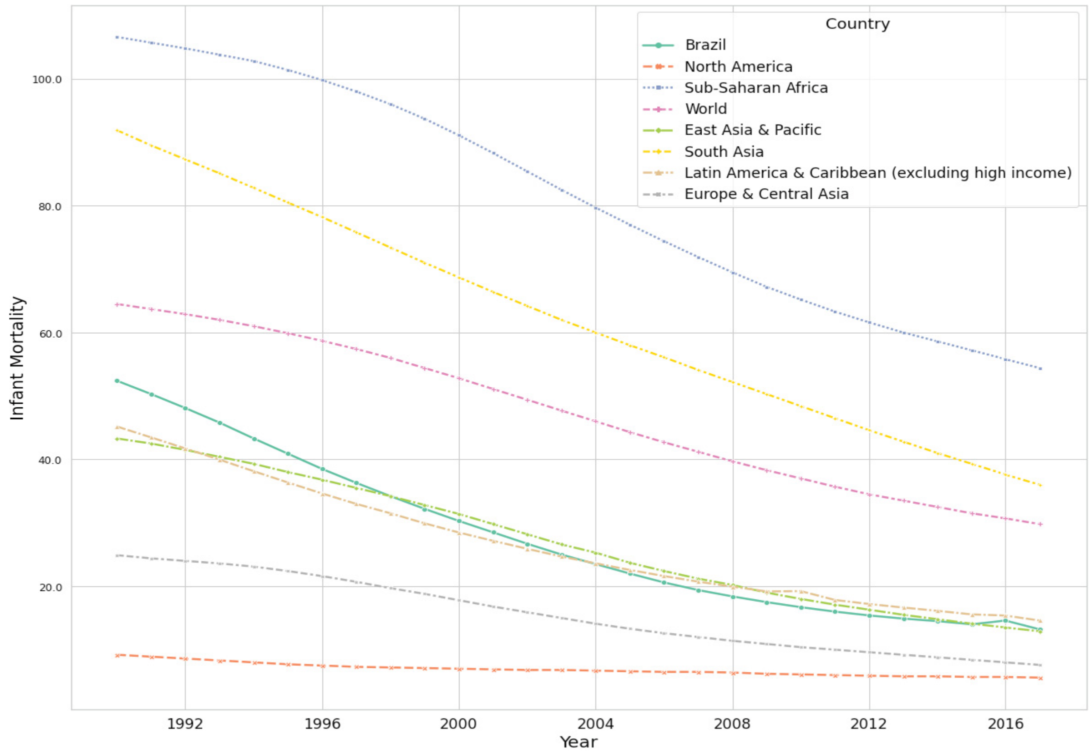
Source: Prepared by the authors using Unified Health System (SUS) data.
Notes: The table describes the features' summary statistics for infants that survived or not the first year of life.

1.4 INFANT MORTALITY IN BRAZIL

Since 1988 with the universalization of healthcare through the Unified Health System (SUS), Brazil has seen an effort to expand health services to its population (PAIM et al., 2011). The nineties were a period of severe economic stress for Brazil, and the main public policy focus was the end of hyperinflation and macroeconomic stabilization. It was not until the 2000s that the economic resources to support large-scale social policies became more available. Indeed, an important driver of infant mortality reduction in the 21st century was the combination of family health expansion policies together with conditional cash transfer programs (GUANAIS, 2015) (RUSSO et al., 2019)

In Brazil, the improvement in social indicators after the 1989's Constitution can be seen in many dimensions but particularly in healthcare (VIELLAS et al., 2014). To get a better perspective on these improvements, figure 1 shows trends in global infant mortality rates using World Bank's infant mortality indicator - specifically, the mortality rate per 1000 live births (BANK, 2021). Following the worldwide pattern of infant mortality reduction, Brazil has reduced its mortality from 47.1 deaths per 1000 births in 1990 to 13.5 in 2015, substantially reducing its gap from the European and North American averages.

FIGURE 1 – Worldwide Trends in Infant Mortality

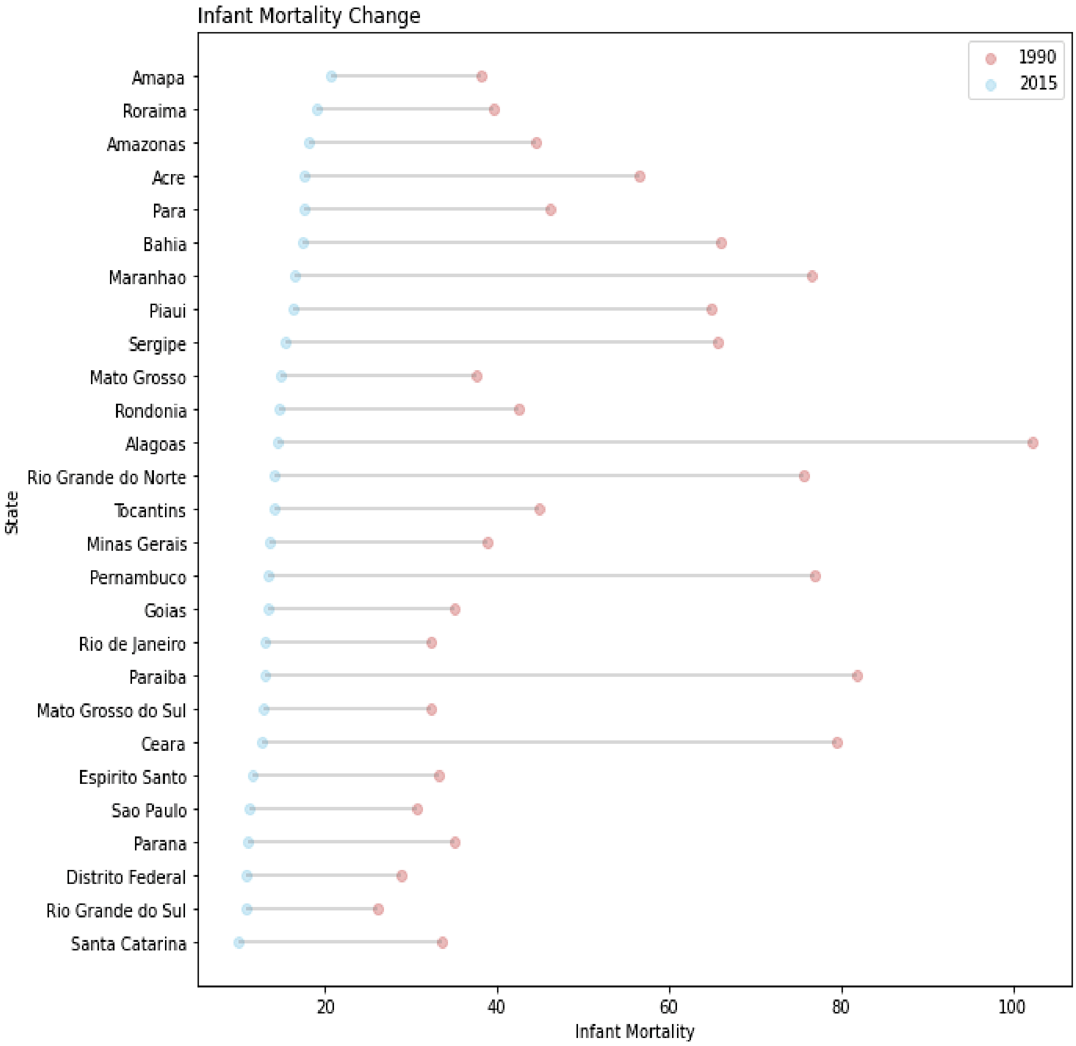


Source: Prepared by the authors based on World Bank data.

Notes: The figure represents the temporal trends of infant mortality in different regions of the world.

From the previous discussion, we have seen that Brazil has had substantial progress in reducing infant mortality. Still, there is substantial inequality *between* states as can be seen in figure 1 that uses regional infant mortality data adapted from Szwarcwald et al. (2020). The worse state, Amapa (located in the northern region), has a 20.8 mortality rate. The best state, Santa Catarina (southern region), has a 9.9 rate. Overall, there is a pattern of richer states in the south and southeast having statistics similar to developed countries. In contrast, poorer states in the north and northern regions are still far from these objectives.

FIGURE 2 – Infant Mortality in Brazillian States

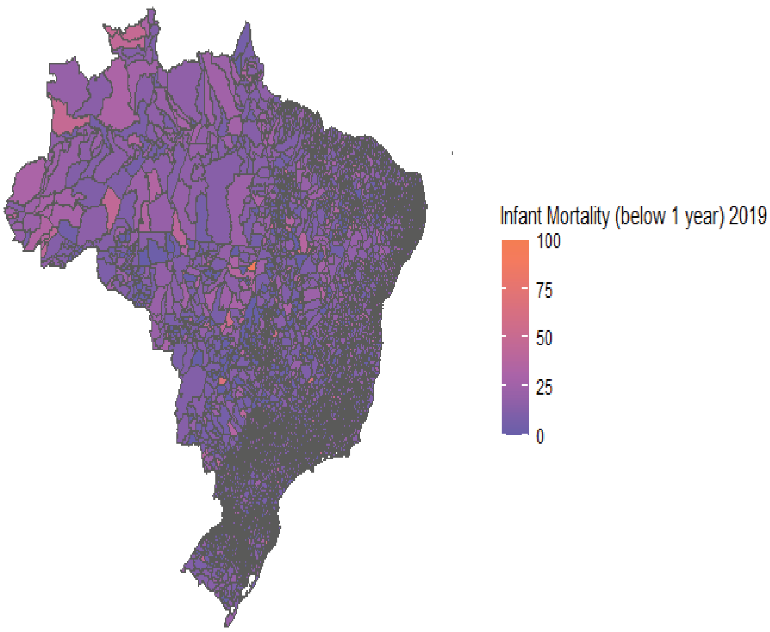


Source: Prepared by the authors based on Szwarcwald et al. (2020) data.

Notes: The figure shows temporal trends of infant mortality in different States of Brazil.

Finally, using data from the livebirth (SINASC) and mortality systems (SIM), figure 3 shows the mortality rate (below one year) per 1000 births in all Brazillian municipalities. We can see that there is still a lot of variation in different municipalities, even within states. The Brazillian unified health system (SUS) has a decentralized institutional framework where municipalities share the responsibility and expenses of providing health services with the federal government. In this way, municipalities with better socioeconomic conditions - such as income, educational attainment, and piped water provision - in general, do better in terms of health outcomes (BUGELLI et al., 2021) (GAMPER-RABINDRAN; KHAN; TIMMINS, 2010).

FIGURE 3 – Infant Mortality in Brazillian Municipalities



Source: Prepared by the authors using Unified Health System (SUS) data.

Notes: The figure represents the map of infant mortality in different municipalities of Brazil. In the color scale, red represents a high infant mortality and purple represents a low infant mortality for the year of 2019.

1.5 METHOD

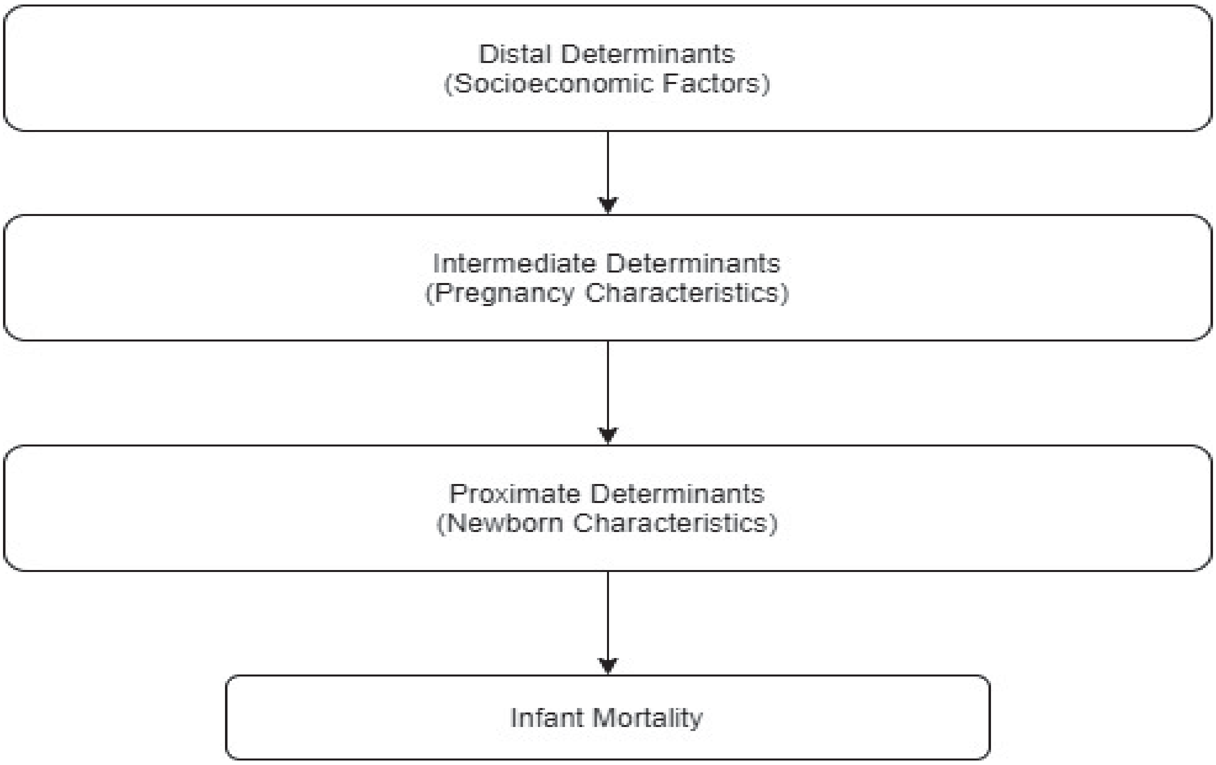
1.5.1 Conceptual Framework

Infant mortality is a complex problem that is better comprehended using a multifaceted framework. That is the critical point of Mosley and Chen (1984) seminal work on infant mortality in developing countries. It defined an analytical framework to integrate the social science, and medical science approaches to understand child survival that influenced a vast number of subsequent studies (ABATE; ANGAW; SHAWENO, 2020).

In Mosley and Chen´s framework, infant mortality is understood as being affected by socioeconomic, biological, and healthcare variables. These factors are organized in a hierarchical structure where different variables can be modeled and labeled according to their importance to the dependent variable (MOSLEY; CHEN, 1984). For instance, socioeconomic variables affect infant mortality mediated by the newborn biological characteristics - which are more important variables.

Studies that use this framework to understand the determinants of infant mortality in Brazil must adapt it to the particularities of the country´s data sources. Therefore, studies in the Brazilian context typically model infant mortality as being influenced by three sets of factors: distal, intermediate, and proximal (GARCIA; FERNANDES; TRAEBERT, 2019)

FIGURE 4 – Mosley and Chen Theoretical Framework



Source: Prepared by the authors based on Mosley and Chen (1984).
Notes: The figure represents a simplified and schematic representation of Mosley and Chen´s seminal infant mortality framework.Distal determinants affect mortality trough its influence in intermediate and proximate ones.

(SOUZA; DUIM; NAMPO, 2019). Figure 4 is a schematic representation. Distal factors are socioeconomic and demographic variables. Intermediate factors are maternal and reproductive variables. Proximal factors are newborn characteristics. Therefore, the pathway through which socioeconomic variables affect infant mortality is through maternal and newborn characteristics.

1.5.2 Empirical Strategy

Our empirical strategy is to use three different survival analysis approaches to understand the determinants of infant mortality in Brazil. The first approach uses a non-parametric estimator of the survival function - the Kaplan-Meier estimator. The second approach is to use the cox regression model- the classical model for survival analysis in statistical inference. Our third approach is to use different machine learning models to estimate the survival function (MONCADA-TORRES et al., 2021) (CHMIEL et al., 2021). The usage of statistical inference methods such as the Cox regression model along with machine learning methods aims to understand the research problem at hand - infant mortality - through the two main approaches of statistical modelling.[3]

[3] Breiman (2001b) argues that there are two cultures in statistical modeling. Roughly speaking, the first culture, called data modeling culture, is the one that dominates the statistical community because the main objective is to interpret the parameters involved in the model; in particular, there is interest in hypothesis testing and confidence intervals for these parameters. Under this approach, testing whether the model's

1.5.2.1 Descriptive Survival Analysis

In survival data, the main component of descriptive analysis is the function of survival: defined as the survival probability of an observation until a specific time t. The survival function is characterized as a ladder function with steps at the observed times of death. The initial procedure of the descriptive phase is to find an estimate for the survival function and then estimate the statistics of interest - particularly the average or median time of survival.

A particular challenge of finding estimators for the survival function is that survival data are characterized by censoring: observations that are incomplete or partially observed in the dataset (GIJBELS, 2010). In an infant mortality context, the typical censored pattern found in datasets is right censoring. For instance, an infant drops out of the study before the end, and therefore their outcome is unknown, or if they die at some unknown point after the end of the study design.

The Kaplan-Meier estimator for the survival function is a common choice in that it can deal with censoring in the data and has desirable statistical properties (KAPLAN; MEIER, 1958). It is a non-parametric estimator with a structured procedure to determine the survival curves: the probability of an individual surviving up to time t is the product of the probability of surviving each of the previous times. A fundamental assumption is that the probability of survival up to time t is estimated considering that the survival until each time is independent of the survival until other times(BLAND; ALTMAN, 1998). From the Kaplan-Meier estimator, it is possible to compare visually different survival curves according to different qualitative variables.

1.5.2.2 Cox Regression Model

The most used model in survival analysis is the Cox regression model (COX, 1972). The key feature of the Cox model is that it assumes that hazards are proportional. Based on this proportionality, it is possible to estimate the effects of covariates on the survival probability without any assumptions regarding the distribution of survival time. No statistical distribution is assumed for the hazard function, only that the covariates act multiplicatively on the hazard.

It is important to assess the suitability of the Cox regression model to the particular modeling problem that we are interested in and the particular dataset at hand. That is, to assess if the covariates and data being used in the model are in accordance with the proportional hazards assumption. The violation of this basic assumption can lead to inconsistencies in the estimation of the model coefficients (O'NEILL, 1986). Model evaluation techniques are based on the Cox regression residuals, and Schoenfeld's residue is the most common technique

assumptions are valid is important. The focus is on inference rather than on prediction. The second culture, called algorithmic modeling culture, is the one that dominates the machine learning community. In this, the main objective is the prediction of new observations. It is not assumed that the model used for the data is correct; it is only used to create good algorithms to predict new observations well. There is often no explicit probabilistic model behind the algorithms used.

(SCHOENFELD, 1982). The hypothesis test is then to check the correlation coefficient between the standardized Schoenfeld residuals and a function of time for each covariate. Correlations close to zero show evidence in favor of the assumption of proportional hazards.

As discussed above, the Cox regression model relies on the assumption of proportional hazards, which in many applications is not reasonable. In addition, there might be non-linearity in the relationship between infant mortality and the covariates that are not straightforward to incorporate in parametric models - the usual alternative to the Cox regression model framework. However, both these challenges can be solved using specific machine learning models.

1.5.2.3 Machine Learning

Statistical learning or machine learning refers to a set of prediction tools designed to understand the available data (FRIEDMAN; HASTIE; TIBSHIRANI, 2001).[4] The function of machine learning algorithms is to discover the relationship between the variables of a system, its inputs, and outputs, from sampled data (CHERKASSKY; MULIER, 2007)[5]

There are usually two types of problems in this literature. The first is supervised learning, where for every training data (or pattern) available, there is a known correct answer. In this case, we say the data is labeled. The second is unsupervised learning, where there is no desired output associated with each pattern, so the data is unlabeled. In this scenario, we want the model to capture, represent or express properties existing in the dataset. All the models used in this research are supervised statistical learning models.

The first step in building good prediction functions to discover the underlying data patterns is to create a criterion to measure the performance of a given prediction function. This is typically done through the mean square error in a regression context. When we measure the performance of an estimator based on its quadratic error, creating a good prediction function is equivalent to finding a good estimator for the regression function (FRIEDMAN; HASTIE; TIBSHIRANI, 2001). Indeed, estimating the regression function is, in this sense, the best way to create a function to predict new observations based on observed covariates - i.e., to learn patterns from data.

Therefore, the purpose of regression methods from a predictive perspective is to provide,

---

4 Following Friedman, Hastie and Tibshirani (2001), let us assume that we observe Y and n different variables: $X_1, X_2, ..., X_n$. Moreover, we suspect that there is some relationship between X and Y. So we can write:

$$Y = f(X) + \epsilon \tag{1.1}$$

f is assumed to be some *fixed* but an unknown function of the variables $X_1, X_2, ..., X_n$, and of ϵ an error term that is independent of X and has zero mean. Thus, f represents systematic information of the set of characteristics that X transmit about the result Y. Statistical learning then (Machine learning) is the set of different approaches and tools used to estimate f.

5 An analogy to machine learning is a doctor progressing in residency: learning rules from the data. Starting with observations at the patient level, the algorithms analyze a large number of variables, looking for combinations that reliably predict outcomes (OBERMEYER; EMANUEL, 2016).

in different contexts, methods that present good estimators of the regression function, that is, low error estimators. Therefore, we want to choose a function within a class of candidates that has good predictive power (low quadratic error).[6] Choosing functions with minimum error can induce a methodological error: learning the parameters of a prediction function and testing them with the same data. This model would have a solid predictive performance in the particular sample where it was trained, but it would generalize poorly with unseen data.[7]

A standard method in statistical learning to obtain a model with a better capacity for generalization consists of dividing the dataset into different subsamples or sets. The training set will be used to adjust the model's parameters, and the validation set will be used to monitor the model's generalizability, which then is put to prove in the test set. An important assumption is that observing model performance against validation data indicates how it will behave when exposed to samples not seen in training. In other words, the validation performance is interpreted as an estimate of the generalization error. Thus, minimizing the error with the validation set is expected to increase the generalizability. Therefore, it is expected that the configuration of the model that leads to the smallest error with the validation set, which is not used for parameter adjustment, has the best possible performance with new samples (GUYON et al., 1997).

A more elaborate technique for splitting the data is called k-fold cross validation (REFAEILZADEH; TANG; LIU, 2009). It consists of dividing the set of samples available for training into k folders and carrying out k training sessions, each considering k -1 folders to adjust the parameters and one folder for validation. Every available sample will appear k -1 times in the training set and one time in the validation set. The k training sets will have a different composition, as will the k validation sets. The model's performance is then taken as the average of the performances in the k validation folders. Since cross-validation gives us a way to infer the quality of generalization of a model, this technique is used to choose values for a model's hyperparameters - settings or configurations of the model that cannot be estimated from data (PROBST; BOULESTEIX; BISCHL, 2019).

The overall description of the machine learning models above is adapted to different supervised learning algorithms to deal with different empirical contexts. In survival analysis,

---

[6] The task of training the model is to find the best θ parameters that best fit the training data x_i and y_i results. To train the model, we need to define the objective function to measure how well the model fits the training data. A characteristic of objective functions is that they consist of two parts: the loss of training and the regularization term:

$$obj(\theta) = L(\theta) + \omega(\theta) \tag{1.2}$$

Where L is the training loss function and ω is a regularization term. The loss function is a measure of the model's predictive power concerning the train set. The regularization term controls the so-called "complexity"of the model, avoiding the problem of overfitting (FRIEDMAN; HASTIE; TIBSHIRANI, 2001).

[7] The second situation is underfitting, where the model was not able to adequately approximate the actual mapping, not even in the data used in training. This can occur because the degree of flexibility of the model is insufficient given the complexity of the mapping to be approximated, or also by convergence problems of the training process (JABBAR; KHAN, 2015).

random survival forests, survival support vector machines, and extreme gradient boosting algorithms are standard choices with desirable statistical properties.

1.5.2.4 Survival Support Vector Machine

Survival Support Vector Machine (SSVM) algorithms are an extension of support vector machines (SVM) that can account for censoring in the data (PÖLSTERL; NAVAB; KATOUZIAN, 2015). The original SVMS was developed for binary classification purposes. It does so, in short, by building a hyperplane as a decision surface that separates the distinct classes in a particular dataset. The 'support vectors' are the data points that have a minimum distance from the separating hyperplane (JAMES et al., 2013). However, not all datasets have linearly separable patterns. For non-linearly separable patterns, therefore, the method uses an appropriate mapping function to make the mapped set linearly separable.[8]

One class of mapping functions that are computationally efficient are called kernels - functions that can project data from lower to higher dimensional spaces. One right choice of a transformation function will result in a higher dimensional feature space that is separable. The statistically robust method to choose a particular kernel is to use cross-validation techniques (JAMES et al., 2013).

Support Vector Machines in the context of survival analysis can be understood in two ways. The first is a ranking problem, where the model aims to understand the accurate ordering of data observations (samples) according to their survival time - it is ranking observations according to their risk. The second is a regression problem, where the model learns to predict the survival time of different observations (PÖLSTERL; NAVAB; KATOUZIAN, 2015). Survival Support Vector Machines uses the kernel function transformation to deal with non-linearities between the variables - a valuable characteristic in predicting survival times.

1.5.2.5 Ensemble of Machine Learning Models

Different machine learning models can solve a particular problem differently, which makes the idea of combining these models in a committee pertinent. The diversity of solutions can make the committee more robust in terms of generalizability (TRESP, 2001). In a particular kind of committee, an ensemble, machines are trained from the available data and some kind of combination of their outputs. The idea of combining machine learning models is based on the idea that diversifying perspectives can bring a better generalization - a model that performs better with unseen data (BISHOP, 2006)

Consider a regression problem where there is a need to approximate an ideal function from a set of data. An ensemble is a set of individual regression models. If we assume that the

[8] More formally, let the input set S be represented by the pairs (x_1,y_1),...,(x_n,y_n), where y_i, i=1,2,...n is the label of each input i. The feature space is a space of higher dimensionality in which the input set S will be mapped, using a function ϕ, in order to obtain a new linearly separable data set S', represented by $(\phi(x_1)$,y$_1)$, ..., $(\phi(x_n)$,y$_n)(JAMESetal., 2013)$.

errors have a null mean and are uncorrelated - an idealized condition - the ensemble will have a mean square error smaller than the mean of the individual errors. One way to approach this condition is to perform a bootstrap aggregation procedure - bagging - to generate the datasets of the individual machine learning models. In this procedure, if there is data in the training set, each set will be composed of samples obtained with replacement (BISHOP, 2006).

Another classic ensemble approach is called boosting. In this case, a sequential training scheme is adopted; the machines are trained in sequence. The training of each machine is based on a dataset in which the data is weighted according to the performance of the previous machines (SCHAPIRE, 1999).

A common ensemble machine learning model that uses bagging is the random forest model, whereas the boosting method is used in the extreme gradient boosting model. Both of these models have been adapted to address survival analysis problems.

1.5.2.6 Random Survival Forest

In the case of ensembles, the idea is that each machine seeks to deal with the task at hand from different perspectives, generating answers that, combined, can lead to a better generalization. The application of this idea in the context of decision trees gives rise to the random forest concept (BREIMAN, 2001a). The extension of this ensemble to the notion of random forest typically involves the construction of trees from subsets of features.

Decision trees configure methods that use a tree-based graphical representation, whose objective is to identify groups of individuals with characteristics of common interest. For this purpose, a recursive method divides the initial sample into subsamples based on observed results of the explanatory variables and their interactions. The tree induction process is started through a sample called a root node divided into subsamples, called child nodes or intermediate nodes. These subsamples, when subdivided, are called parent nodes, as they generate child nodes. When a subsample can no longer be subdivided according to some stopping criteria, it is called an end node or leaf node. This process is called recursive because each subsample generates new subsamples (SONG; YING, 2015).

The Random Forest method combines the idea of bagging and the random selection of explanatory variables in the tree induction process. In this case, a data set of N samples is sampled (with replacement), generating M "new" sets, which, in turn, are used to build M trees. The responses from these trees are combined to generate the ensemble's output. This selection is a drawing made at each tree node, randomly selecting some candidate variables to divide this node. Using this technique, different sets of variables may appear at different levels in each tree. With this, the technique becomes more sensitive to interactions between variables, in addition to resulting in decorrelated trees, due to the random drawing of candidate variables to divide the node made in each partition (BREIMAN, 2001c).

Random Survival Forests are a particular kind of ensemble tree suited to analyzing survival data. The rules for splitting the trees in the model use censoring and survival information (ISHWARAN et al., 2008). To create the samples and subsamples - the parent and child nodes - in growing the decision tree, the random survival forest algorithm allows different splitting criteria. The most intuitive and used one is the log-rank criteria: where the splits follow the difference in survival times between groups, as measured by the log-rank test (BOU-HAMAD; LAROCQUE; BEN-AMEUR, 2011). A key advantage of Random Survival Forests is that interactions between variables and non-linearities do not come automatically out of parametric model survival models. In contrast, random survival forests can deal naturally with these challenges because of their decision tree structure (ISHWARAN et al., 2008).

1.5.2.7 Extreme Gradient Boosting

The Extreme Gradient Boosting (XGBoost) algorithm is an ensemble model. As we described earlier, they are methods that use a combination of results from weak predictors - called base learners - to produce a better predictive model. Weak predictors - or weak learners - are models that, when used individually, have an accuracy that is marginally better than a random guess (CHEN; GUESTRIN, 2016).

It is important to highlight that the random forest and gradient boosting models are similar. They both are based on weak predictors to make the final prediction, however, the random forest model uses an average of predictions from weak learners, whereas the gradient boosting model uses the boosting method. In the Boosting technique, each weak classifier is trained with a set of data, sequentially and adaptively, where a base model depends on the previous ones, and in the end, they are combined in a deterministic way. It builds the model in stages and generalizes them, allowing the optimization of an arbitrary differentiable loss function (SCHAPIRE, 1999).

The choice of the loss function is particularly relevant for survival analysis because it can impact our assumptions regarding the hazard distribution. One can choose the partial likelihood loss function based on the Cox regression model class or the loss function weighted by the logarithm of survival time of the accelerated failure time model class. The first is based on the proportional hazard assumption, whereas the latter allows for time-varying hazards (WEI, 1992).

More precisely, the accelerated failure time model takes the following form in the Gradient Boosting context:

$$lnY = \tau(x) + \delta Z \tag{1.3}$$

in this equation, lnY is the logarithm of the survival time, $\tau(x)$ is the result of a decision tree ensemble, given our vector of controls x. Delta is a scaling parameter. Z is a random variable with some definite probability distribution that we will assume is normal. The Gradient Boosting model maximizes the log-likelihood of Y using the decision tree output $\tau(x)$ (BARNWAL; CHO; HOCKING, 2020).

Given a loss criteria, then, the objective of the Gradient Boosting algorithm is to create a chain of weak models, where each one aims to minimize the error of the previous model. These interactions are repeated a certain number of times, seeking to minimize the residual generated by weak models, that is, until the distance between the predicted value and the actual value is as small as possible. The final model is the sum of the fits of all weak models. The adjustments of each weak model are multiplied by a value called the learning rate - which controls the complexity of the model (the propensity of overfitting). This value is intended to determine the impact of each tree on the final model - the smaller the value, the smaller the contribution of each tree (CHEN; GUESTRIN, 2016).

1.5.2.8 Model Evaluation and Interpretation

It is typical to use the Concordance index (C index)- a measure of the correlation between the model predictions and the data - to assess the model results from the machine learning models (HEAGERTY; ZHENG, 2005). The C index has a close relationship with the AUC: *the area under the receiver operating characteristic (ROC) curve.*[9]. It is particularly suited to survival analysis because it handles censoring in the data.

Understanding the predictions of a particular method is a crucial part of choosing modeling strategies. Decision tree models such as Random forests and XGboost allow the implementation of feature (variable) importance methods. In brief, feature importance can be measured by how much an accuracy metric changes when a feature (variable) is not used. In survival analysis settings, it is typical to rank the variables according to their impact on the C-index as a measure of their importance in the model prediction. However, most tree-based algorithms only provide the global aggregate importance of a particular variable but do not provide the direction of the impact - positive or negative (ROGERS; GUNN, 2005).

A method to interpret the predictions of machine learning models is to use Shapley values. They are originally a concept of game theory. In coalition game theory, a group of players comes together to create some value. For instance, one can think of a group of people coming together to form a company to generate profit. The Shapley value is a method of distributing the profit fairly among players based on their contributions. More generally, the Shapley value is the average marginal contribution of a characteristic value across all possible coalitions (SHAPLEY; ROTH et al., 1988).

Shapley values have been adapted to interpretable machine learning into the SHAP

[9] The area under the ROC curve (AUC) measures the capacity of a given test to assess whether a particular criteria is present or not. The standard interpretation is: 1.0 AUC represents perfect discrimination capacity, whereas 0.5 represents a test with no capacity. (HOO; CANDLISH; TEARE, 2017)

framework: a unified approach to interpret model's predictions. The objective of the SHAP framework is to interpret the prediction of any instance of the machine learning model by computing the contribution of each feature to the prediction (LUNDBERG; LEE, 2017)

1.5.2.9 Sampling Strategy

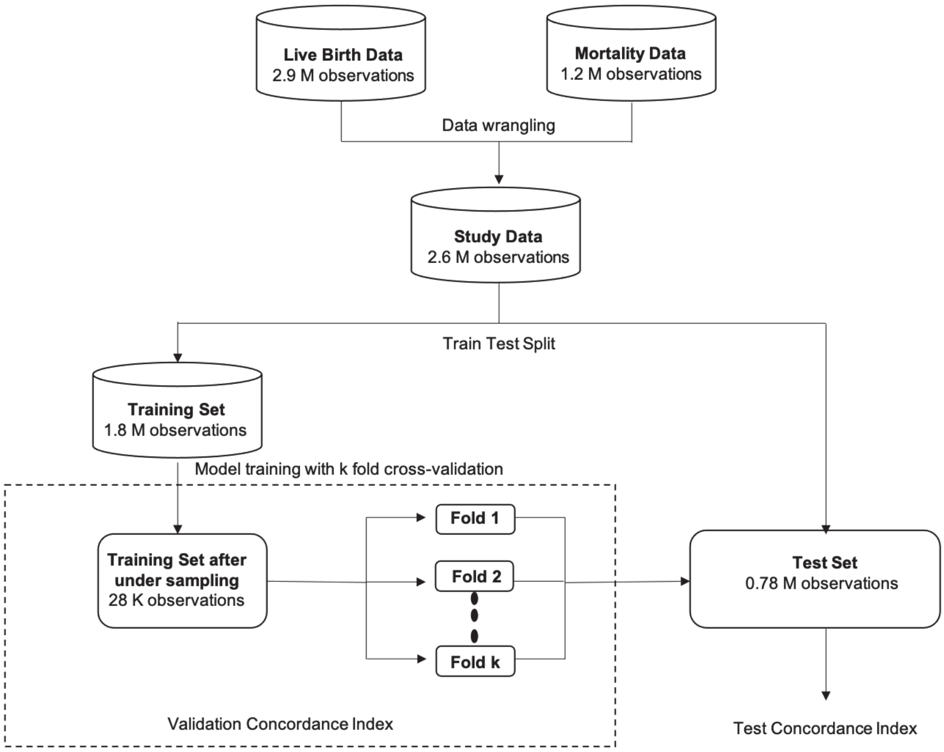
Our dataset consists of 2.65 million birth and 20.5 thousand deaths. The event that we are trying to comprehend, infant mortality, is present in a small proportion of the data. That is, the dataset is heavily unbalanced. Unbalanced data can be defined by the small incidence of a category within a dataset (minority class) compared to the other (majority class). In most cases, this means that we have much information about the most incident category and less about the minority one (BRANCO; TORGO; RIBEIRO, 2016).

Unbalanced data can cause problems in machine learning models and their predictions. Traditional ML algorithms will favor the unbalanced class heavily because of their objective functions (CIESLAK; CHAWLA, 2008). For instance, when 99% of births do not result in deaths, the most straightforward prediction is to infer that every newborn will survive. That will result in a 99% accuracy metric. However, the model will be useless to identify newborns that have a lesser likelihood of survival.

One way to remove the bias caused by the difference in the proportion of the categories is to alter the amount of data that the machine learning models effectively use. A typical method is undersampling, reducing the number of observations of the majority class to reduce the difference between the categories. The result is a dataset that has a similar number of observations between the classes (DRUMMOND; HOLTE et al., 2003).

There are different undersampling strategies fit for different purposes and challenges. The simplest strategy is random undersampling, which consists of randomly removing data from the majority class - the drawback is the inevitable loss of information. However, the method can be efficient in different contexts (ESTABROOKS; JO; JAPKOWICZ, 2004) (CHAWLA; JAPKOWICZ; KOTCZ, 2004).Another common method is using distance criteria to evaluate which observations in the imbalanced class should be added to the training set. For instance, the nearest neighbor algorithm is used to define a relative distance between observations, and then the data is undersampled to preserve the information structure revealed by the algorithm. (MANI; ZHANG, 2003).

FIGURE 5 – Sampling Strategy



Source: Prepared by the authors.

Notes: The figure shows the sampling strategy used by the authors.

An important precaution is to split the data between train and test samples before applying an undersampling strategy. If the undersampling algorithm is applied to the test set as well, we will have a data leakage or train-test communication problem. That is, information from the test set will leak to the training set, which will probably overestimate the model's predictive performance. For the results to be robust, a machine learning model cannot be evaluated in the same sample that it is trained (FILHO; BATISTA; SANTOS, 2021).

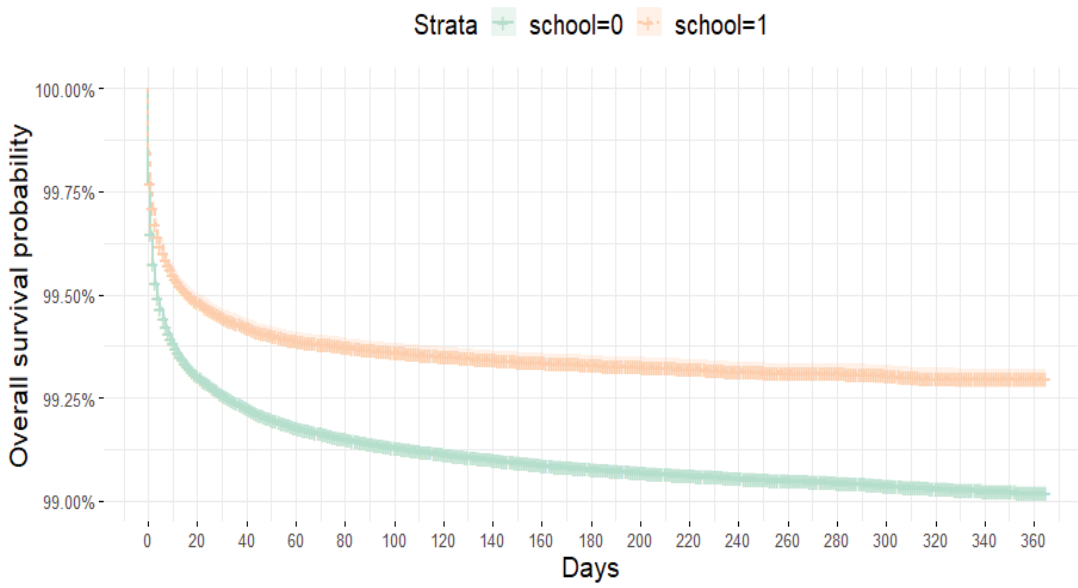
After the sample splitting, we first balance the train test using a random undersampling algorithm to summarize our procedure. We also tested the distance criteria method, but the results were similar, so we chose random undersampling for computational efficiency. The models are estimated and optimized using k-fold cross-validation. Afterward, the model performance is measured in the test set. Figure 5 describes the sampling strategy beginning with the original microdata coming from the Brazilian Unique Health System (SUS).

1.6 RESULTS

1.6.1 Descriptive Survival Analysis

The typical first step in descriptive survival analysis is the analysis of Kaplan Meyer survival curves for different groups. We highlighted one specific explanatory variable for each factor group in Mosley and Chen (1984) infant mortality framework. Namely, schooling from the distal group, mother having a C-section from the intermediate, and infant sex from the proximate factor group. Figure 6 shows the difference in survival probabilities by the schooling level dummy - a distal factor. Mothers with higher schooling (in red) are associated with an increase in the infant´s probability of survival.

FIGURE 6 – Kaplan Meyer: School

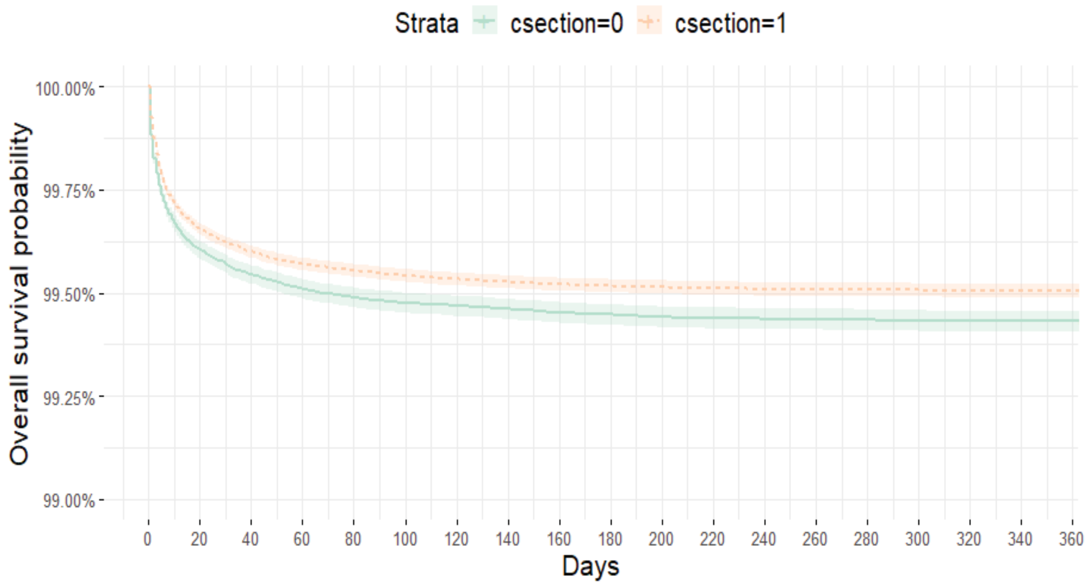


Source: Prepared by the authors using Unified Health System (SUS) data.

Notes: The figure shows the Kaplan Meyer survival curves for different levels of mother's schooling. The blue line represents the curve for mothers that have finished college, and the red line represents mothers who did not.

Figure 7 shows the survival difference between the group of mothers how have done a c-section and the group who have not. The delivery type is typically considered an intermediate factor in Mosley and Chen (1984) framework. The Klapan Meyer curves suggest that babies born from c-section (in red) are associated with a higher probability of survival. This descriptive result is at odds with the literature on the health impacts of c-section in Brazil, which find negative results for the infant´s health (PAIXAO et al., 2021b). The association between c-section and the probability of survival will be discussed more thoroughly in in the machine learning model results section, where we will see how non-linearity is an important element in the relationship.

FIGURE 7 – Kaplan Meyer: C-Section

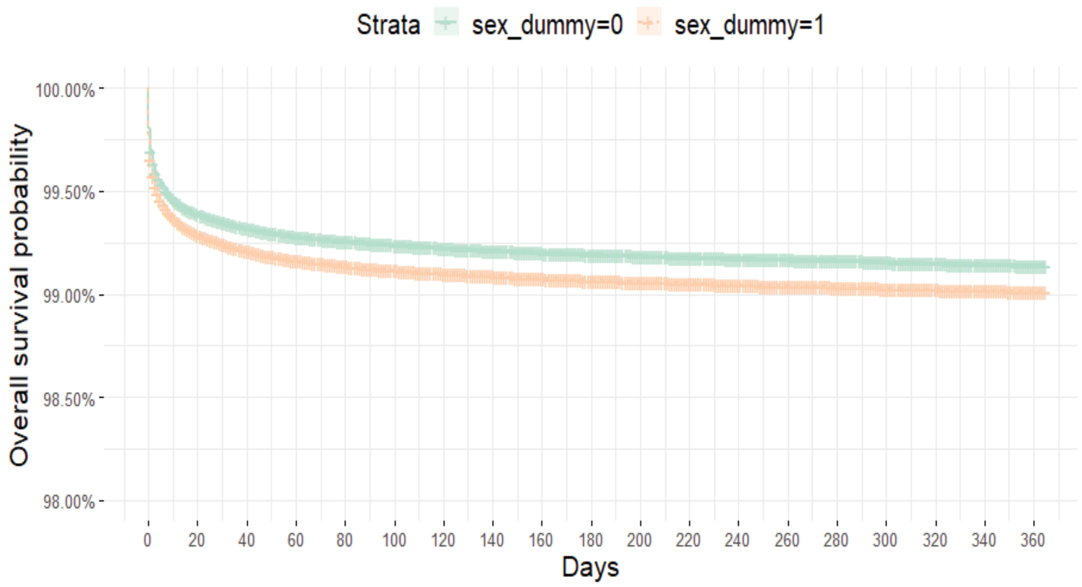


Source: Prepared by the authors using Unified Health System (SUS) data.

Notes: The figure shows the Kaplan Meyer survival curves for type of delivery. The blue line represents babies born from vaginal births, and the red line represents babies born from c-section.

Figure 8 shows the survival differences by the infant´s sex - a proximate factor. Male infants have a slightly lower survival probabilities than female ones - a typical result in infant survival analysis (NASEJJE; MWAMBI, 2017).

FIGURE 8 – Kaplan Meyer: Sex



Source: Prepared by the authors using Unified Health System (SUS) data.

Notes: The figure shows the Kaplan Meyer survival curves for babies gender. The blue line represents female babies, and the red line represents male babies.

1.6.2 Cox Regression

Table 4 shows the results for the Cox regression model. The regression table is structured to highlight the distal, intermediate, and proximal factors. In the distal factors, schooling has a statistically significant adverse impact (-0.135) on the hazard (increases the probability of survival), whereas living in a frontier city positively impacts the hazard (reduces the probability of survival). Being married (marital status) has a negative impact (-0.073) but less statistical significance.

In the intermediate factors, parity (0.075) and number of dead children (0.045) have a statistically positive impact on the hazard, whereas having a c-section (-0.258), having induced labor (-0.273) or assisted labor (-0.197) , and having more prenatal visits (-0.013) reduce the hazard. In the proximal factors, low APGAR score (2.162), low weight (2.521), having a genetic anomaly (2.068), and being born a man (0.228) have a statistically significant positive impact in the hazard ratio.

TABLE 4 – Cox Regression Results

Variable	**Coefficient (se)**	**Pr(>|z|)**
Distal Factors		
Mother Age	-0.001 (0.002)	0.636
Father Age	0.008 (0.002)	0.396
Capital Residency	-0.027 (0.037)	0.450
Schooling	-0.135*** (0.035)	0.0001
Marital Status	-0.073* (0.029)	0.013
Border Residency	0.234*** (0.044)	1.30e-07
Intermediate Factors		
C Section	-0.258*** (0.031)	<2e-16
Prenatal Visits	-0.013*** (0.002)	5.45e-07
Parity	0.075* (0.031)	0.015
Birthplace Dummy	-0.214 (0.175)	0.222
Dead Children	0.045** (0.017)	0.009
Induced Labor	-0.273*** (0.046)	4.87e-09
Assisted Labor	-0197* (0.082)	0.017
Proximal Factors		
Low APGAR1	2.162*** (0.029)	<2e-16
Low Weight	2.521*** (0.031)	<2e-16
Anomaly	2.068*** (0.129)	<2e-16
Fetus Presentation	24.47584	3
Race Dummy	0.054 (0.027)	0.051
Sex Dummy	0.228*** (0.27)	<2e-16
Number of Observations	2666742	
Number of events	20310	
Concordance Index	0.896 (se = 0.003)	
Likelihood ratio test on 25 df	21338	<2e-16
Wald test = 28077 on 25 df	28077	<2e-16
Score (logrank) test on 25 df	67422	<2e-16

Source: Prepared by the author.
Notes: The table shows the results for the cox regression models. Standard errors are in parenthesis. *p<0.1; **p<0.05; ***p<0.01

Table 5 shows the results for the proportional hazards assumption using the Schoenfeld residuals test. Mother age, father age, the schooling dummy, marital status dummy, and living in border dummy are not statistically significant, whereas living in the capital is. The birthplace dummy, number of dead children, and the assisted labor dummy are not statistically significant, whereas having a c-section, the number of prenatal visits, parity, and having induced labor is. Finally, having a low Apgar score, low weight, genetic anomaly, the type of fetus presentation are statistically significant. Overall, distal factors satisfy the proportional hazards assumption, whereas intermediate and proximal factors do not.

TABLE 5 – Proportional Hazards Assumption Test

	Schoefelnd Test		
Variable	**Chi Squared**	**Degrees of Freedom**	**p-value**
Distal Factors			
Mother Age	0.550	1	0.458
Father Age	0.005	1	0.939
Capital Residency	13.544***	1	<0.001
Schooling	0.286	1	0.592
Marital Status	0.738	1	0.390
Border Residency	0.011	1	0.915
Intermediate Factors			
C Section	5.621**	1	0.017
Prenatal Visits	5.336**	1	0.020
Parity	14.271***	1	<0.001
Birth Place Dummy	0.011	1	0.914
Number of dead children	0.033	1	0.855
Induced Labor	3.426*	1	0.064
Assisted Labor	3.496	4	0.478
Proximal Factors			
Low APGAR1	316.379***	1	<0.001
Low Weight	66.618***	1	<0.001
Genetic Anomaly	36.031***	2	<0.001
Fetus Presentation	24.475***	3	<0.001
Race Dummy	0.027	1	0.869
Sex Dummy	2.172***	1	<0.001
GLOBAL	476.448***	25	<0.001

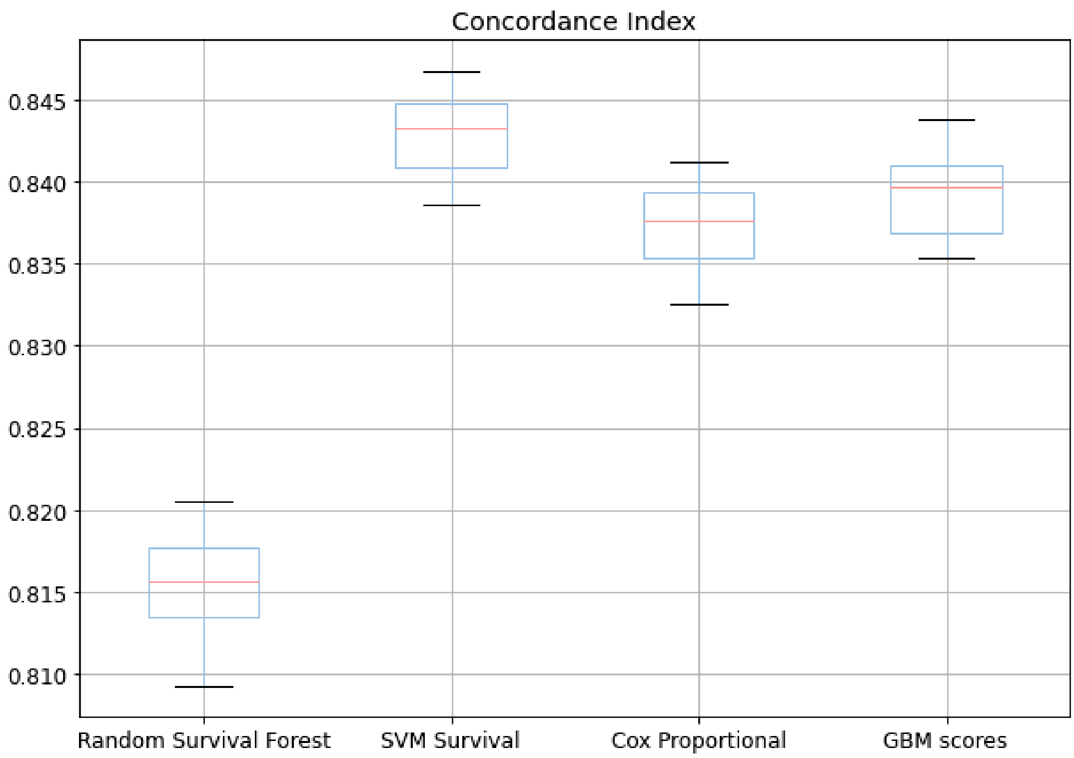
Source: Prepared by the authors using Unified Health System (SUS) data.
Note: The table describes the results of the Schoefelnd Test for the proportional Hazard assumption. The null hypothesis is that the hazards ratios are proportional.
*p<0.1; **p<0.05; ***p<0.01

The Schoefeldn test indicates that a significant set of covariates do not satisfy the proportional hazards assumption in our sample. Estimating a Cox proportional model with non-proportional hazards has consequences. There can be an overestimation of risks if hazards are increasing and an underestimation if hazards are converging (SCHEMPER, 1992). The interpretation of model coefficients can be misleading in this setting. If the objective is the efficient prediction of survival probabilities, this empirical challenge can be tackled using machine learning models that are robust to non-proportional hazards.

1.6.3 Machine Learning Models

Figure 9 describes the machine learning model results. It shows the average predictive performance of the different ML models in the test set as measured by the concordance index score. The Cox proportional hazards model (0.837) is a benchmark to assess the other models. The model with the best predictive performance is the SVM Survival (0.843), followed by the Gradient Boosting (0.839). The Random Survival Forest (0.815) has the worse predictive performance. Only the SVM Survival and the Gradient Boosting have a higher predictive performance than the Cox model.

FIGURE 9 – Machine Learning Model Performance



Source: Prepared by the authors using Unified Health System (SUS) data.

Notes: The figure shows the average c-index for the machine learning models using the k-fold cross-validation method.The Random Survival Forest model has a mean c-index of 0.815, (standard deviation: 0.003). The SVM Survival model has a mean c-index of 0.843 (standard deviation: 0.002). The Cox Proportional Hazards model has a mean c-index of 0.837 (standard deviation: 0.002). The Extreme Gradient Boosting has a mean c-index of 0.839 (standard deviation: 0.002.)

Table 6 describes the hyper-parameters that were used to estimate the models. The model parameters were optimized using the randomized search cross-validation method [10]. We provide a brief description of each. In the SVM model, the alpha parameter (0.113) controls

[10] Randomized search implements a random search of parameters, where each configuration is sampled from a distribution of possible parameter values (HACKELING, 2017)

the model's regularization: higher α value increases the model bias but reduces its variance [11]. The chosen kernel transformation was the radial basis function (RBF)[12] . Although, in general, the Cox proportional hazards model is not considered a machine learning algorithm, there are ways to estimate it in such a way as to be used as a benchmark to the other models. So instead of using the full sample to estimate the model - as in the statistical inference method - the sample is divided between train and test sets and a regularization parameter is set to control the model complexity. Therefore, the only hyper-parameter in the Cox proportional hazards model is the alpha parameter (0.232) which controls the model´s regularization - analogous to the SVM alpha.

TABLE 6 – Hyper-parameter Tuning

Model	**Parameter**	**Chosen Value**
Random Survival Forests	Number of Trees	100
	Maximum Depth	6
	Minimum Samples Leaf	1
Gradient Boosted Models	Number of Trees	50
	Maximum Depth	1
	Learning Rate	0.5
Survival Support Vector Machines	Kernel	RBF
	alpha	0.113
Cox Proportional Hazards	alpha	0.232

Source: Prepared by the authors.
Note: The table describes the parameters for each model. Models were parameterized using a randomized search of different parameter settings to maximize the models predictive performance.

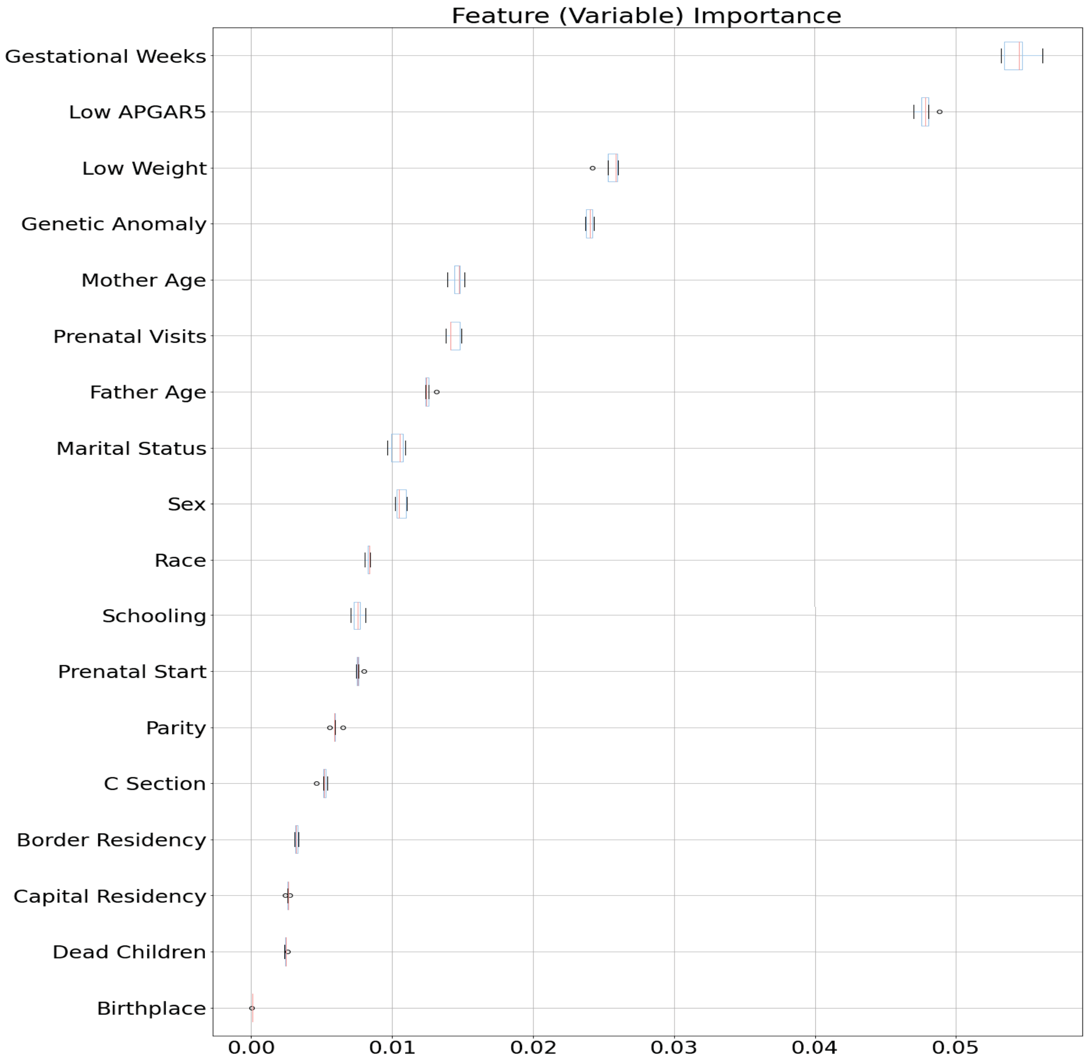
The decision tree models' important hyperparameters are the maximum and minimum depth of the tree, which determine the number of decisions that the tree will make. We would expect that the deeper the tree is, the more decisions it has to make and the more perfect its training would be against our tests. However, this does not happen because, at very large depths, the tree becomes so perfect for the training data that it fails the test data (overfitting). For the Random Survival Forests (RSF), the maximum depth is six and the minimum depth is 1, whereas, for the XGBOOST, the maximum and minimum depth is 1. The total number of trees is 100 for the RSF and 50 for the XGBOOST. Finally, the XGBOOST model has a learning rate parameter (0.5) which controls the complexity of the model (SOMMER; SARIGIANNIS; PARNELL, 2019).

[11] *Bias* is the inability of a model to capture the true relationship between variables and the object to be predicted. The model is not learning. On the other hand, if there is a very small bias, the model is so adjusted to the training data that when used with different data, it makes many mistakes. The model is overfitted. *Variance* is the sensitivity of a model to being used with datasets other than training. If the model is very sensitive to the training data, it identifies the relationship between them so well that it will be very inaccurate when faced with different data. *regularization* is a method that seeks to penalize the complexity of models, reducing their variance (TIAN; ZHANG, 2022)

[12] Kernel functions are intended to project feature vectors into a high-dimensional feature space for the classification of problems that lie in non-linearly separable spaces. The model is then able to classify the output variables into different categories (MUSAVI et al., 1992).

To interpret the model predictions, figure 10 shows the feature (variable) importance for the Random Survival Forest model. Gestational weeks is the most relevant predictor, followed by low APGAR5, low weight, genetic anomaly, mother age and the number of prenatal visits, Similar to the Cox regression results, the most important variables to the probability of survival are proximate factors. The distal and intermediate factors have less importance on the model´s predictions.

FIGURE 10 – Variable Importance



Source: Prepared by the authors using Unified Health System (SUS) data.

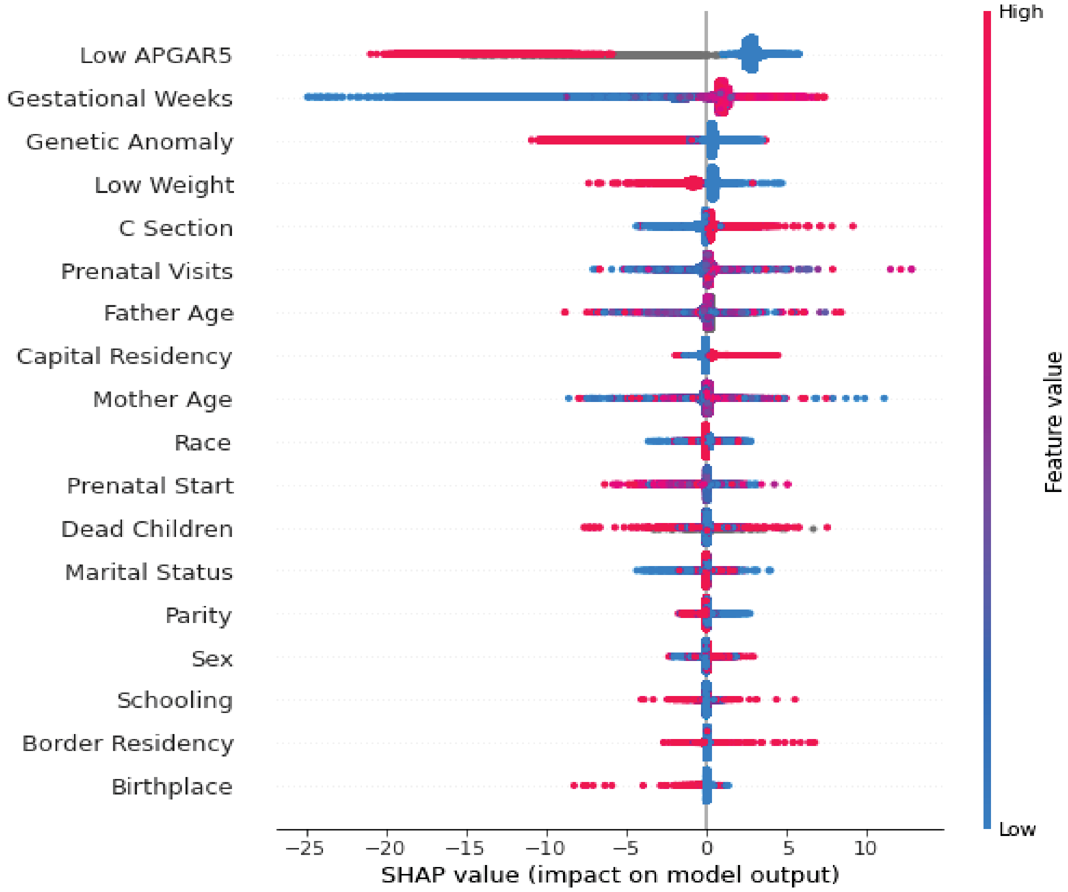
Notes: The figure represents a feature importance of the Random Forest Survival model.he variables are ordered along the y axis based on their importance. That is, the higher the variable is on the y axis, the more important it is for the model prediction

Feature importance can be misleading if there is no proper understanding of the relationship between explanatory variables. Indeed, we highlight that one important characteristic of the feature importance algorithm is its sensitivity to the correlation between explanatory variables (features). In particular, the correlation between features should be considered when interpreting these results. [13]. This is not a significant concern if the model is for prediction purposes only. It will not impact the model performance in assessing the likelihood of survival. However, the contribution of each variable to the prediction is hard to assess.

Also, as we discussed in the empirical strategy section, feature importance algorithms do not show in which direction each variable impacts the model, only the relative importance to the prediction. Figure 11 then shows the SHAP values for all explanatory variables in the XGBOOST model. Low APGAR, too few gestational weeks ,low weight, and genetic anomaly are the most important predictors of mortality, and they impact the hazard ratio in a negative way. They are proximate factors. Having a c-section and the number of prenatal care visits - intermediate factors - are important to the model output. Finally, distal factors such as schooling and marital status have less impact on the model predictions. This pattern is in harmony with the results in the Cox regression model.

[13] This challenge is similar to the multicollinearity problem in econometrics. There, if two independent variables are strongly correlated, the estimates of the coefficients of the model parameters can become insignificant since each one presupposes, by definition, the variation in Y given the variation in X. A high correlation will cause both variables to move together, and it will be hard to disentangle the particular effect of each one (ALIN, 2010). The difference in a Random Forest's feature importance setting is that there is no statistical significance, and what is entangled is the contribution of each variable to the model predictions.

FIGURE 11 – Model Interpretation: Shapley Values



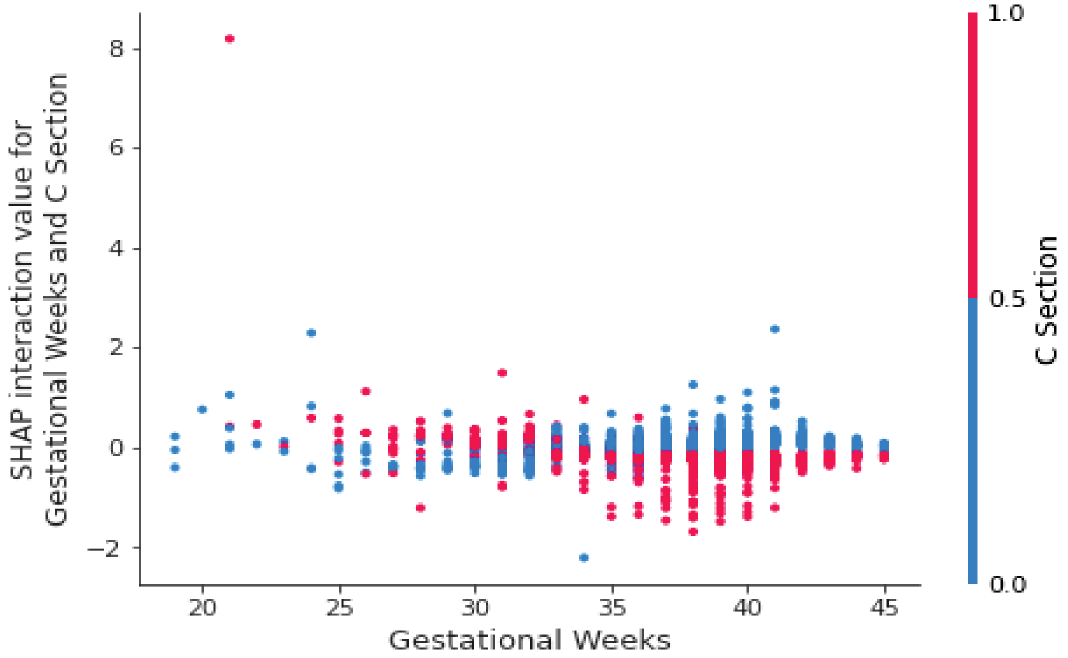
Source: Prepared by the authors using Unified Health System (SUS) data.

Notes: Summary plots for SHAP values. For each variable, the points correspond to different observations. The SHAP value is the impact of the specific variable (feature) for that specific observation. This corresponds to the survival probability relative across observations, where a higher SHAP value has a higher survival probability relative to a lower SHAP value. The variables are ordered along the y axis based on their importance, given by the average of their SHAP values. The higher the variable is on the y axis, the more important it is for the model prediction.

Figure 14 in the appendix shows the SHAP interaction values between the features in our models. In the main diagonal are the mean effects of each variable on the model prediction. In the off diagonals, there is the interaction between variables. Most variables can influence the model prediction differently when interacting with others. In particular, some variables have interactions that are worth highlighting. They reveal insights that are interesting to the infant mortality problem.

For instance, figure 12 shows the dependence plot between gestational weeks and having a c-section. Babies with fewer gestational weeks had a higher likelihood of survival if they had a c-section - the SHAP value is higher. As gestational weeks increase, the relationship inverts. Babies with more than 35 weeks have a higher likelihood of survival if they had a normal birth - having done a c-section decrease the SHAP value. There is a non-linear interaction between c-section and gestational weeks.

FIGURE 12 – Dependence Plot: Gestational Weeks vs C Section



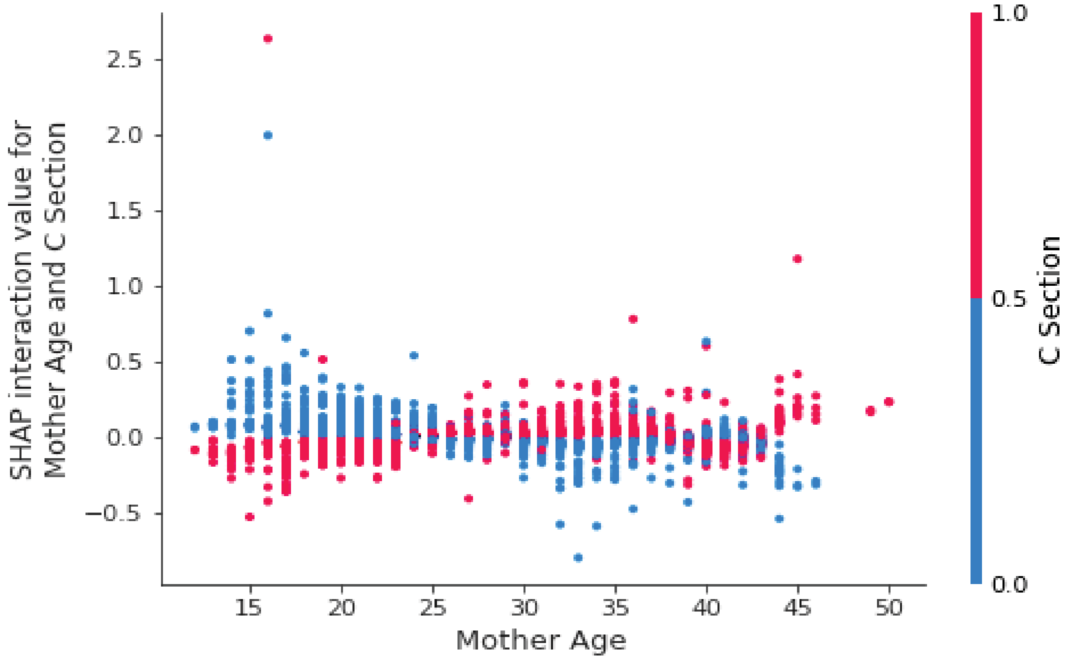
Source: Prepared by the authors using the Unified Health System (SUS) data.

Notes: The figure shows the SHAP feature dependence plot of the XGB model for the interaction between gestational weeks and the low weight dummy. The plot shows how the two variables affect the probability of survival non-linearly.

Figure 13 shows the SHAP dependence plot between the mother's age and having done a c-section. The graphic also shows a non-linear relationship between the two variables. Being a teenage mother - between 10 and 20 years - and having done a c-section decreases the probability of survival (the SHAP value is negative). As age increases between 20 and 39 years, there is no straightforward relationship between c-sections and infant survival. Increasing the age even further - more than 40 years - having done a c-section now increases the probability of survival (SHAP value is positive).

We highlight that this result should be interpreted with care. Teenage pregnancies tend to have more adverse outcomes for children (OGAWA et al., 2019). There is, in general, more chance of preterm delivery, low birth weight, and fetal distress (BAŞ et al., 2020). A sampling selection effect could cause the relationship that the model is showing. The teenage mother group has higher risk pregnancies and therefore has a higher chance of doing c-sections (YUSSIF et al., 2017). The model then indicates a relation between doing a c-section and a decrease in the probability of survival. However, without properly accounting for pregnancy riskiness in the model, there is no way to discern if the decrease in the survival rate comes from having a c-section or an underlying omitted factor.

FIGURE 13 – Dependence Plot: Mother Age vs C Section



Source: Prepared by the authors using the Unified Health System (SUS) data.

Notes: The figure shows the SHAP feature dependence plot of the XGB model for the interaction between age and c-section dummy. The plot shows how the two variables affect the probability of survival non-linearly.

The SHAP framework illustrates how interpretable machine learning can shed light on non-linear relationships omitted in mean effects. Explanatory variables can affect the outcome variable differently depending on their positions in the data distribution. The mean effect in some settings might cause an important loss of information. For instance, in the Cox proportional model, the mean effect of c-sections is to increase the survival ratio. However, the SHAP dependence plots show that this omits important contexts where c-sections reduce the survival ratio.

1.7 DISCUSSION

1.7.1 Cox Regression

The Cox regression model results have a discernible pattern where intermediate and proximal variables have a more significant and substantial impact on mortality than distal factors. Low weight, low APGAR, and genetic anomaly are the most important drivers of the hazard ratio, whereas schooling and marital status have a much lesser impact. This result is in accordance with the theoretical framework of Mosley and Chen (1984) and the subsequent empirical studies done using it.

A correctly specified model that utilizes the whole set of intermediate and proximate factors will, in consequence, tend to have distal factors that are not statistically significant.

That is because the proximate determinants capture all the variance in the model. However, it is unrealistic to assume that the actual variables coming from health data sets can measure all proximate aspects of infant mortality correctly. That is why including socioeconomic factors are important, and there is, in general, statistical significance in variables of this group (HILL, 2003).

The Cox regression results are consistent with prior works in the same literature for Brazil. For instance, low birth weight being an important risk factor for mortality (RISSO; NASCIMENTO, 2010; CARDOSO et al., 2013; PAIXAO et al., 2021a). Also, the probability of survival being negatively affected by the mother´s having fewer prenatal visits and schooling (PINHEIRO; PERES; D'ORSI, 2010; GARCIA; FERNANDES; TRAEBERT, 2019).

1.7.2 Machine Learning Models

The Random Survival Forests (RSFs), Survival Support Vector Machines (SSVMs) and the Extreme Gradient Boosting (XGBOOST) machine learning models achieve a good prediction performance since they all have a high concordance index, that is the models can predict efficiently if an infant will survive the first year of life. However, only the SSVMs (C-index: 0.843) and the XGBOOST (C-index: 0.839) models have a slightly better predictive performance than the Cox proportional model used as a benchmark (C-index: 0.837). The RSFs have a worse predictive performance ((C-index: 0.815).

On the one hand, we can argue that this finding contributes to an emerging literature that shows the good performance of survival analysis using machine learning methods that are robust to non-proportional hazards (MONCADA-TORRES et al., 2021) (CHMIEL et al., 2021). On the other hand, the good performance of the Cox model, in the presence of non proportional hazards, can be interpreted as a signal of its strength and its substitution for more complex and computationally intensive models should be done with care. Particularly when there is a loss of interpretability when using 'black-box' machine learning models.

1.7.3 Prediction Frameworks

A common challenge in machine learning applications is assessing the true relationship between explanatory variables and the model output (GILPIN et al., 2018). On the one hand, machine learning models are very efficient for prediction purposes. For instance, as the SHAP results show, a model identifying that baby's from teenage mothers (who had a c-section) have a higher risk of death is an efficient prediction. On the other hand, the relationship between features and outcomes are not guaranteed to be stable across different machine learning modelling strategies. That is, our feature importance and SHAP results should be taken with a grain of salt and they should not be interpreted as a causal effect that can be generalized to other settings.

We argue that any predictive model in healthcare should be tested and discussed with subject matter specialists before being put into production. Machine learning is best at prediction problems and can be used as a tool in policy-making contexts where prediction is the crucial aspect (KLEINBERG et al., 2015). However, if the estimation of causal effects is the main focus, an empirical strategy that answers counterfactual questions is required (ATHEY, 2017). The danger is to infer causality from a model that is designed to be efficient in prediction and not in answering causal questions. If predictive models are put into production without having these caveats in mind, there is a danger that may cause more harm than good - particularly in terms of inducing wrong or unfair decisions in healthcare (MEHRABI et al., 2021).

1.8 FINAL REMARKS AND POLICY IMPLICATIONS

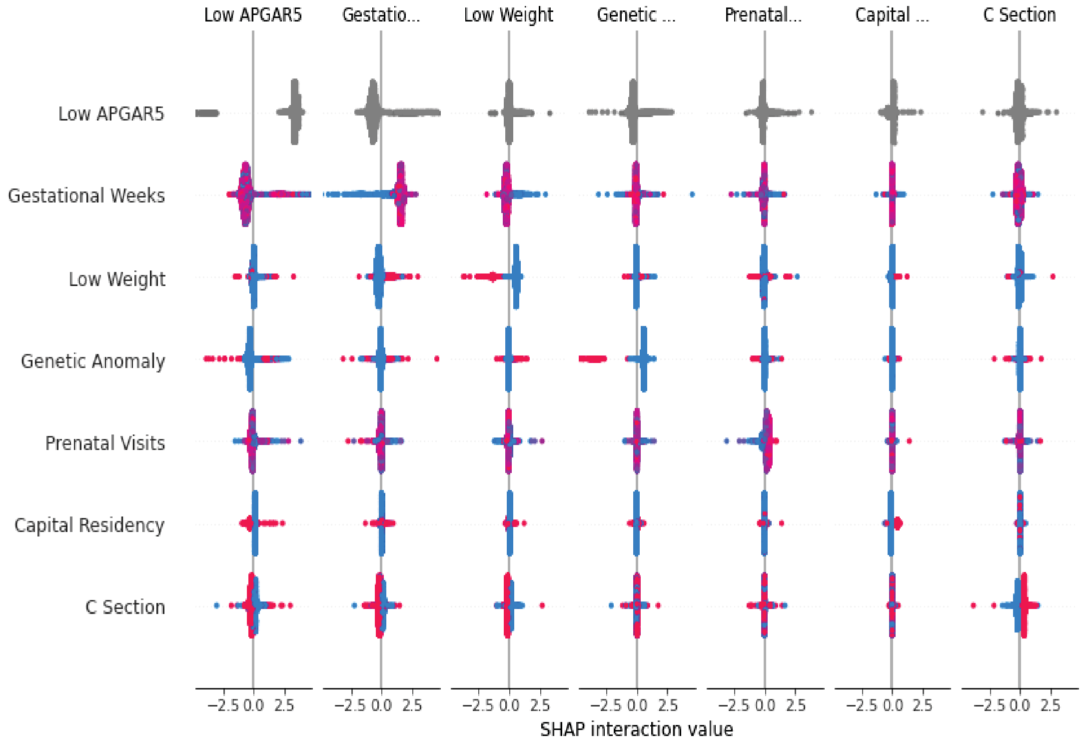
Infant mortality is a serious public health challenge worldwide because, despite the global decrease in its rates, it is still a stark reality in several developing countries. In the last decades, Brazil has markedly improved newborn health conditions, which has greatly reduced its infant mortality ratios. Nevertheless, there is still substantial room for improvement, particularly in less developed regions. Therefore, reducing infant mortality is still a major challenge for Brazilian policymakers and society as a whole.

This paper contributes by using survival analysis with machine learning models that are efficient in predicting infants at risk of death, as well as the risk factors associated with mortality. Specifically, Random Survival Forests, Survival Support Vector Machines, and Extreme Gradient Boosting models can achieve a concordance index higher than 0.8 in the task of predicting mortality in the first year of life. Furthermore, using the SHAP framework, we provide evidence that variables such as gestational weeks, low weight, and having a cesarean section interact non-linearly in affecting mortality. To our knowledge, this is the first research using survival analysis and machine learning for infant mortality in Brazil.

These findings have policy implications for Brazil since identifying newborns that have a high risk of death at the moment of their birth can be a valuable input in health policy. Naturally, this requires an accurate prediction of survival probabilities, a task that machine learning models are efficient. Model predictions - in particular interpretable machine learning models - can be incorporated into a policy framework that can help mitigate infant mortality by being proactive in assessing risks. Finally, future researchers should integrate machine learning strategies with causal analysis frameworks to create more transparent and robust models that can tackle health problems more efficiently.

1.9 APPENDIX 1

FIGURE 14 – SHAP Interaction Values



Source: Prepared by the authors using the Unified Health System (SUS) data.

Notes: The figure shows the SHAP interaction values. The main effect of each variable in the model result is shown in the main diagonal. The interaction effects between variables are shown by the intersection of each pair of variables outside the main diagonal.

2 PRENATAL CARE AND BIRTH WEIGHT IN BRAZIL: REGIONAL AND SOCIOECONOMIC INEQUALITIES

2.1 ABSTRACT

This study examined the impact of prenatal care on the newborn´s weight, using data from a sample of 5 million births between 2015 and 2017 from the Brazillian Unique Health System (SUS). A fixed-effects instrumental variable was used to measure the prenatal care impact on birth weight, addressing the endogeneity and heterogeneity bias inherent in this setting. Our findings are that each prenatal care visit has a positive mean effect of 70 grams in the birthweight and each delayed month in prenatal care has a negative mean effect of 75 grams. To assess the robustness of these findings to different models, a propensity score matching framework was used to assess the impact of inadequate prenatal care on the probability of low birth weight (<2500g). An inadequate number of prenatal care visits increase the odds of low weight (<2500g) [odds ratio (OR): 2.715] and very low birth weight (<1500g) [odds ratio (OR): 8.220] for newborns in the sample. Given these findings, we argue that public policies aiming to increase prenatal care attendance in Brazil can improve newborns' health outcomes, particularly in poorer regions and socially fragile groups.

Keywords: Prenatal care; Newborn health; Instrumental variable; Propensity Score Matching; Health economics; Brazil

2.2 INTRODUCTION

There is an increasing awareness that health outcomes in newborns are crucial. This has extended to the economics literature, where a rich body of evidence is showing that early childhood is fundamental for the complete development of individuals (CAMPBELL et al., 2014). In the context of early childhood, the first days of life are highly relevant for a child's healthy development. Health decisions about childbirth can have substantial and lasting consequences in the lives of newborns. In one particular dimension, low birth weight (LBW), there is evidence that differences in health at birth can have long-term impacts. For instance, LBW has been associated with lower levels of income in adulthood, and in particular, in Brazil, there is evidence of a substantial relationship between LBW and infant mortality (CURRIE, 2011; CARRILLO; FERES, 2017). According to the World Health Organization, prenatal care - provided to pregnant women to ensure adequate health conditions for mothers and babies - can diminish perinatal mortality through the identification of pregnancy risk factors and adjustments in nutrition and care for the mothers (ORGANIZATION et al., 2016). An adequate amount of prenatal care is important to mitigate newborns' health problems, particularly LBW.

A case-control study in Campinas, an industrial city in the state of São Paulo, finds that low education, delayed prenatal care start, and few prenatal visits are important risk factors for LBW (COUTINHO et al., 2009). Another case-control research in Botucatu, São Paulo, finds a significant relationship between prenatal care inadequacy and LBW, even though an increase in prenatal care does not necessarily translate into higher birth weight (FONSECA et al., 2014). A cross-sectional epidemiological study for the municipality of São Paulo identifies mother age, marital status, low schooling, and inadequate prenatal care as risk factors for LBW (MENDES et al., 2015). For Teresina, a city in the Brazilian northeast region, there is evidence that mothers with inadequate prenatal care have a greater chance of having children with LBW and premature birth, which is a pathway that can cause newborn death (GONZAGA et al., 2016).

A study using three birth cohorts in São Luis, another city in the Brazilian northeast region, shows that substantial improvements in schooling and prenatal care are not necessarily associated with a reduction in the odds of LBW (VELOSO et al., 2014). Another cohorts study (1979-1994) in Ribeirão Preto, a city in the Brazilian state of São Paulo, shows that although the proportion of mothers with adequate prenatal care increased substantially, the proportion of children with LBW did not decrease. A risk factor analysis shows that inadequate prenatal care was associated with low birth weight in the 1979 cohort but not in the 1994 one (GOLDANI et al., 2004). A retrospective cohort of 8.8 million births in Brazil finds that the most significant risk factors for LBW have a black mother, low education, lower number of prenatal visits, and being primiparous (FALCÃO et al., 2020). However, none of these studies can account for selection bias or confounding due to unobservable variables.

Therefore, a relevant methodological challenge is to estimate the relationship between prenatal care and birth weight with the least possible bias. A literature review shows that studies have found a positive association between prenatal care and birth weight, but identification issues must be appropriately addressed. Indeed, this relationship suffers from endogeneity - as mothers with high risk might self-select into having more prenatal care - and therefore calls for an identification strategy to mitigate this problem (SILVEIRA; SANTOS, 2004). Ordinary least squares (OLS) estimation of this relationship produces bimodal residuals, which indicate the need to adjust for normal and complicated pregnancies by using finite mixture models and a two-stage least squares strategy (CONWAY; DEB, 2005). Estimating a birth weight production function through OLS produces biased results, which can be mitigated with a two-stage least squares model using marital status as an instrument for prenatal care (JEWELL; TRIUNFO, 2006). Using a bus strike as an instrument, a study finds evidence of the importance of prenatal care on early pregnancy but little evidence for late pregnancies (EVANS; LIEN, 2005).

To address the issue of selection bias, Wehby et al (2009) uses a well specified quantile regression framework to measure the impact of prenatal care and birth weight in Brazil. For Kenya, Awiti (2014) deals with the endogeneity and sample selection bias by using a multi-level model estimation and finds a positive association between prenatal care and birth weight.

The author also argues that there is a public policy implication of reducing the distance from pregnant mothers' residencies to prenatal care clinics. Finally, a research on a nationally representative level sample in Mexico, using an instrumental variable model, also finds a positive effect of prenatal care in newborn's health results (GONZALEZ; KUMAR, 2018).

This paper's contribution is to assess the relationship between prenatal care and birth weight in Brazil, by using an instrumental variable method and a large health micro-data. The data sample is 5M births between 2015-2017 from the Unique Health System (SUS). To our knowledge, this is the first work in Brazil to assess the prenatal care effect on birth weight using a large sample of nationally representative data.

The modelling strategy is a fixed-effects two stage least squares (2SLS) instrumental variable framework to deal with endogeneity and account for heterogeneity in hospitals and municipalities. Additionally, to assess the results robustness to different modelling strategies, we also perform a propensity score matching method to deal with selection bias and then a logit model in the matched data to assess the impact of inadequate prenatal on the odds of LBW among newborns.

Findings for the instrumental variable strategy are that prenatal care visits have a positive mean effect of 70.26 grams on newborn weight. In contrast, delayed prenatal care start has a negative mean effect of -75.23 grams. Both prenatal visits and prenatal delay are robust to the problem of weak instruments. Also, the impact of delayed prenatal care is consistent with using one or many instruments. The propensity score matching strategy findings are that inadequate prenatal care visits increase the odds of low birth weight [odds ratio (OR): 2.715; IR: 2.594-2.843] and very low birth weight in newborns [odds ratio (OR): 8.220; IR: 7.195-9.431]. All together, we argue that these results form one of the most comprehensive empirical studies of the impact of prenatal care on birth weight in Brazil.

2.3 DATA

This paper's main sources of data are from the Brazillian Ministry of Health´s (MS). First is the Live Birth Information System (SINASC), whose data collection instrument is a standardized document – the Live Birth Declaration (DN) identifier. It includes information on mother and newborn characteristics and is filled by the health professionals at the time of the baby's birth. Secondly, The National Registry of Health Establishments (CNES) is an information system that registers information regarding the healthcare workforce and installed capacity of Brazilian health establishments. All public or private health units in Brazil need to provide information using their CNES identifiers.

Every birth has a unique DN number coming from the SINASC and a unique CNES number from the health unit where the birth happened. The unique identifiers of both datasets allow the linkage between them into a single dataset with newborn, mother, and hospital characteristics. We end up with a dataset between 2015 and 2017 of live births in all Brazilian municipalities amounting to a sample size of 7.3 million births In addition, we were also able to include in our sample information of the place of birth, such as Gross Domestic Product (GDP) and population per municipality information, coming from the Brazilian Institute of Geography and Statistics (IBGE) and the share of health insured population per municipality coming from the Supplementary Health Agency (ANS). Table 7 below shows the description of the variables used in the present study.

TABLE 7 – Variables Description

Variable Name	**Definition**
Birth Weight	Baby's Birth weight in grams
Prenatal Visits	Number of pre-natal care visits
Month of Start	Month which prenatal care started
Birth Place	Hospital or not
C-Section	C-section Birth or not
Mother Age	Mother´s age in years
Marital Status	Mother´s Marital Status
Mother's Education	Mother´s Education in levels
Gestational Weeks	Number of Gestational Weeks
Parity	Type of Pregnancy (Unique; Double; Triple)
Sex	Infant Sex
Race	Race/Ethnicity
Genetic Anomaly	Genetic Anomaly
Assisted Birth	Health Professional who assisted the labor
Previous Gestations	Number of Previous Gestations
Induced labor	Dummy Variable that indicates whether labor was induced or not
Fetus Presentation	Fetus position before Labor
GDP per Capita	Gross Domestic Product per capita by municipality
Insured Population	Share of insured population per municipality

Source: Prepared by the authors using Unified Health System (SUS) data.
Notes: The table describes the variable name that we adopted in our estimations, the definition of each variable based on SUS data, the variable type and the corresponding data source from the SUS.

The outcome variable for this study is Birth Weight (grams) and the Prenatal variables are Prenatal Visits and Month of Start. Birth Place is a dummy variable that indicates whether the birth took place in the hospital or not. C-Section is also a dummy variable, and indicates if it was a C-Section or a vaginal birth. Gestational Weeks is the number of gestational weeks at the time of birth. Parity indicates the number of babies in the pregnancy. Mother's Education is a categoric variable indicating the mother's level of schooling. Race is a categorical variable indicating the mother's race. Fetus presentation is a categoric variable indicating if the fetus's position is cephalic, pelvic, or transverse. Induced labor is a dummy variable that is one of the labor was induced and zero otherwise. Assisted labor is a categorical variable indicating which health professional assisted the mother during labor.

Table 8 shows key descriptive statistics for selected numeric and categorical variables in our sample. The average number of prenatal visits in the sample is 7.79, higher than seven, the recommended number by the Ministry of Health (MS). Also, the average month of prenatal start is 2.59, also within the recommended criteria of starting prenatal before the first gestational trimester (BRASIL; SAÚDE, 2012). The average birth weight is 3193.83 grams with 552.178 grams of standard deviation. Father´s (31.0 years) are in general older than Mothers (26.4 years). The proportion of children born with the genetic anomaly is 0.84% of all births. Hospitals are responsible for 98% of all births, with 1.46% happening in other locations. Most women (67.2%) in the sample were not married at the moment of birth. Only 19.05% of them had a college degree.

TABLE 8 – Descriptive Statistics

Variable Name	**Definition**	**Mean (SD)**
Prenatal Visits	Number of prenatal Care Visits	7.79 (2.72)
Month of Start	Month of prenatal start	2.59 (1.47)
Birth Weight	Infant's weight in grams	3193.83 (552.178)
Father Age	Father's age in years	31.03 (7.73)
Mother Age	Mother's age in years	26.41 (6.69)
Variable Name	**Category**	**Total (Proportion)**
Genetic Anomaly		
	Infant without genetic anomaly	7262015 (99.16)
	Infant with genetic anomaly	61194 (0.84)
Birth Place		
	Hospital	7216419 (98.54)
	Other	106790 (1.46)
Mother's marital status		
	Married	2402567 (32.80)
	Not married	4920642 (67.20)
Mother's race		
	White	2420988 (33.06)
	Otherwise	4902221 (66.94)
Mother's education		
	College	1395369 (19.05)
	High School	4364909 (59.6)
	Primary	1266168 (17.3)
	Less than 3 years	171772 (2.3)
Observations	7323209	

Source: Prepared by the authors using Unified Health System (SUS) data.
Notes: The table describes the variables' summary statistics for births in our sample of data between the years of 2015-2017.

2.4 CONTEXT

There was no structured public health system in Brazil before the constitution of 1988, which established the Unified Health System (SUS) and defined health as a citizen´s right, assigning to the state the responsibility of its provision (CASTRO et al., 2019). The

TABLE 9 – Proportion of Prenatal Care by Region

		Proportion of Visits			Total number of Pregnancies
Region	**None**	**1 to 3**	**4 to 6**	**7 or more**	
North	0.04%	13.02%	36.05%	50.87%	815008
Northeast	0.03%	7.61%	30.31%	62.03%	2199981
Southeast	0.03%	4.12%	19.00%	76.84%	2702139
South	0.03%	3.53%	16.03%	80.39%	929303
Mid-West	0.05%	5.66%	24.03%	70.24%	676778
Total	0.03%	6.23%	24.38%	69.34%	7323209

Source: Prepared by the authors using Unified Health System (SUS) data.
Notes: The table describes the proportion of prenatal visits in Brazillian regions by different categories between the years of 2015-2017.

SUS system works in a decentralized way where the 5570 Brazilian municipalities have the responsibility of delivery and management of health care services - the financing burden is divided between the federal government and the municipalities (PAIM et al., 2011). In parallel, there has been the development of a private healthcare system in the country. Brazil´s health system then is *segmented* with a decentralized public health provider - used mainly by the poorer parts of the population - and private health care providers - used mainly by the wealthier parts of the population through private insurance [1] (CASTRO et al., 2019).

Prenatal care services are provided in the context of this segmented health system - through the SUS and the private sector - and are regulated by the Brazilian health authorities. Specifically, the Brazilian Ministry of Health guidelines for prenatal care is of at least six visits - which include vaccines, laboratory tests, supplementation, and medical treatment in the event of complications (LEAL et al., 2020). The timing of the visits also matters; the policy directive from the Ministry of Health recommends one appointment in the first gestational trimester, two in the second, and three in the last (BRASIL; SAÚDE, 2012).

Figure 16 in the appendix shows the evolution of prenatal care visits in Brazil since 2000. In 2018, almost all pregnant women in Brazil did at least one prenatal visit, but those that do the recommended amount are around 80 percent. That indeed is a substantial improvement over the 2000 figure of less than 50 percent of all pregnant women doing the recommended amount. However, as noted by Leal et al (2020) regional inequalities and inadequate prenatal care are a concern in Brazil. Indeed, substantial disparities still exist as indicated by table 9 showing the proportion of prenatal care visits by region. Half of all pregnant women in the northern part of Brazil still do not do the recommended amount of prenatal care visits, whereas more than 80 percent of them do in the South. The general picture is that in the wealthier parts of Brazil - the south and southeast regions - the outlook is much more favorable than the poorer parts - the north and northeastern regions.

In addition to regional inequalities, previous studies have shown that the recommended

[1] For complex and expensive treatments, typically, private sector patients switch to the public Unified Health System (SUS) (CASTRO et al., 2019)

proportion is lower in young women with little education and low-income (VIELLAS et al., 2014). A study using a 2012-2013 survey analysis from the Ministry of Health has concluded that just 15 percent of the women in the sample had had adequate prenatal care, including all laboratory tests, supplements, and number and timing of visits. The proportion of adequate care was higher for older women in the southeast region of Brazil and municipalities with a high human development index (TOMASI et al., 2017).

2.5 METHOD

2.5.1 Identification Strategy

The relationship between the infant´s birthweight and the mother´s prenatal care is potentially endogeneous. One particular cause is the presence of unobservable variables that influence both the demand for prenatal care by the mother and the infant´s birthweight. That is, there might be an adverse selection effect into prenatal by mothers with higher risk pregnancies (WEHBY et al., 2009). For example, mothers that have an underlying health condition might expect a riskier pregnancy and therefore demand more prenatal care (JEWELL; TRIUNFO, 2006). In this context, the estimation of a standard ordinary least squares model to assess the impact of prenatal care on birthweight will probably result in biased estimates. Also, Brazil is a continental country with a rich diversity of cultures and ethnic groups in its regions. As discussed in the context section, it has substantial regional inequality in health outcomes (LEAL et al., 2020). This way, there is also differences in health risks stemming from unobserved heterogeneity in Brazilian regions.

Another potential source of variation in health results is differences in hospitals. There are potentially multiple mechanisms that are at play in hospital heterogeneity. For instance, different management practices and organizational structures can influence the quality of care, and the performance of hospitals (ALI; SALEHNEJAD; MANSUR, 2018). Because of learning by doing and productive specialization, high-volume hospitals can have better health results than low-volume ones (AVDIC; LUNDBORG; VIKSTRÖM, 2019). Institutional differences caused by ownership type - such as not-for-profit, public, or privately held - may influence treatment choices and, therefore, patient outcomes (BAYINDIR, 2012). Finally, the market structure and competitive environment faced by a particular hospital can impact the quality of care that it provides (GAYNOR; LAUDICELLA; PROPPER, 2012).

The objective is to estimate the relationship between prenatal care and the newborn weight with the least possible bias. The need to establish a causal relationship between variables in the context of observational data suggests causal inference methods specific to this problem. Indeed, this leads to the question of the proper identification of regression coefficients and the associated identification strategy used to achieve the goal of assessing a causal relationship between prenatal care and birth weight. The chosen econometric technique to deal with these challenges is an instrumental variable approach, more specifically, a two-stage least squares model (2SLS) with fixed effects to deal with unobserved heterogeneity from hospital and municipalities characteristics.

In general, instrumental variable strategies typically require three basic conditions. First, the instrumental variable must be associated with the endogenous variable - relevance condition. Second, the instrument should only affect the dependent variable through its influence in the exclusion restriction of the endogeneous variable. Third, unmeasured confounders should not be present between the instrument and the dependent variable - exchangeability (LABRECQUE; SWANSON, 2018). The instrument relevance condition can be tested directly from the data through an F test; however, the exclusion restriction condition and the exchangeability cannot be completely verified. [2] Therefore, in this paper´s context, candidates for instruments should be correlated with prenatal care and excluded from the newborn weight equation. That is, they should impact birth weight only through their influence in prenatal care (AWITI, 2014). [3]

Table 10 presents the instrumental variables used in this work. First, a dummy variable for marital status (0 for married and 1 for single and divorced) is used as the main instrument: the idea is that marriage increases the likelihood of planned pregnancies, investment in newborn health endowments, and prenatal care but does not affect the newborn´s birth weight directly citetodd2006impact. The second instrument is the number of maternity hospitals per hundred thousand people in the municipality of residence. As discussed in the context section, Brazil has a decentralized health system, and there is a substantial variation in health infrastructure between municipalities. Including the number of hospitals per hundred thousand people indicates the general accessibility of healthcare services for each municipality (WEHBY et al., 2009). An adequate supply of maternity hospitals is not a given for all municipalities in Brazil; therefore, a variable controlling for this can be a source of identification for the model.

TABLE 10 – Instrumental Variables

Instruments	Definition	Reference
Marital Status	Married or Not	(JEWELL; TRIUNFO, 2006)
Number of Maternity Hospitals	Number of maternity hospitals per hundred thousand people	(WEHBY et al., 2009)
Distance	Distance between the city of birth and city of mother residence	(AWITI, 2014)
Insured Population	Percentage of insured population in the municipality	(WANG; TEMSAH; MALLICK, 2017)

Source: Prepared by the authors.
Notes: The table describes the instrumental variables used in this paper as well as their source in the literature.

2 When the model is overidentified, that is, when there are more instruments than endogenous regressors, it is typical to run an overidentification test to assess whether there are invalid instruments (HAHN; HAM; MOON, 2011)

3 As Frick and Lantz (1996) one can also think in a *structure-process-outcome* framework. The outcome is the newborn birth weight. The process is related to biological characteristics of fetal development and gestational age that directly influence birth weight. This process is influenced by variables such as the mother´s age, parity, smoking behavior, nutrition, vitamin intake, and so on. Finally, the structure is underlying characteristics that affect the mother´s choice and behavior. Typical structure variables are marital status, education, health insurance, and income, which indirectly affect birth weight through their effect on process variables.

The third instrument is a distance variable. Ideally, the distance between the mother´s home and the hospital, but that is not available. Therefore, a proxy is used: the distance between the municipality of birth and the municipality where the mother resides. The rationale is that shorter distances are an incentive that increases prenatal care and are not directly related to health outcomes (WEHBY; ULLRICH; XIE, 2012) (AWITI, 2014). Distance is particularly relevant considering the extension and heterogeneity of the Brazilian territory - a mother in the Amazon region, for instance, may have to travel hundreds of kilometers to reach a hospital. The percentage of the insured population in the municipality is also used as an instrument for prenatal care. Health insurance increases the utilization of medical services in general and prenatal care in particular [4] (CURRIE; GRUBER, 1996) (WANG; TEMSAH; MALLICK, 2017).

The instrumental variables at the municipal level (distance, number of maternity hospitals and percentage of insured population) impact the sample size in an important way. The linked dataset described in the data section has 7.3M observations. After creating the municipal instrumental variables there is a loss of 2M observations coming from cities without maternity hospitals. Therefore, the final dataset that is used in the estimations has 5.04M birth observations, with mother and municipal characteristics.

2.5.2 Econometric Model

The chosen econometric model is the fixed effects two stage least squares (FE-2SLS) estimator because it is robust to correlation between unobservables, instruments, and explanatory variables and has the advantage of making no assumption of error distributions and no specification requirements for reduced form equations of endogenous variables (SEMYKINA; WOOLDRIDGE, 2010). The baseline 2SLS econometric model is then specified as follows:

First Stage:

$$Prenatal\ Care = \alpha_1 + \beta_4 Instruments + \beta_5 X + \beta_6 time + \beta_7 \gamma_m + \epsilon \tag{2.1}$$

where prenatal care is a measure of prenatal care - the two measures are the number of visits and month of start; Instruments is a vector of instruments including the marital status of the mother, distance to the hospital, percentage of the insured population, and the number of hospitals per capita; X is a vector of confounding variables, time is the year of birth dummies, γ_m is hospital and municipality fixed effects and ϵ is the error term.

Second Stage:

$$Birth\ Weight = \alpha_2 + \beta_7 Prenatal\ Care + \beta_8 X + \beta_9 time + \phi_m + \mu \tag{2.2}$$

In the second stage regression, birth weight is the newborn´s weight in grams, time is the time fixed effects, ϕ is the hospital and municipal fixed effects, and μ is the error term.

[4] A potential specification concern is using instruments that are in the municipality level together with municipality fixed effects. This is allowed because fixed effects account for unobserved heterogeneity between groups whereas we are controlling for *observed* heterogeneity when we include municipality level variables

Afterward, subsample regressions explore in more detail the heterogeneity in prenatal care impact by various characteristics such as county income level, hospital institutional setting, and mother´s ethnicity. The method estimates an FE-IV model for the sub-populations of interest and then assesses the change in the relevant coefficients.

2.6 RESULTS

2.6.1 Instrumental Variable

Table 11 concerns the first stage regressions for both fixed effects models.[5] The number of obstetric hospitals beds per capita, marital status, and the insured rate are all statistically significant. Indeed, the F statistic for instrument relevance is above 10 for the two models indicating that the chosen instruments are relevant for prenatal care measures. Weak instruments are a problem because the 2SLS estimator becomes biased in small samples, and exclusion restrictions violations are magnified. On the other hand, relevant instruments are more robust to bias and to failures in complying with the exclusion restriction (ISAIAH et al., 2018).

---

[5] Most models in this section are using robust standard errors to mitigate heterokedasticity. In particular, given the covariace matrix $VAR(\hat{\beta}) = (X^T X)^-1X^T\omega X(X^T X)^-1$ we are estimating the covariance matrix diagonal ω trough $\hat{\omega} = diag(w_1, ..., w_n)$ using the following estimator: HC3: $w_i = \frac{\hat{u}_i^2}{(1-h_i^2)}$. For additional details, please refer to Zeileis (2004).

TABLE 11 – First Stage Regression

	Dependent variable:	
	Prenatal Visits	Month of Start
	(1)	(2)
Distance	−0.0004*** (0.00003)	−0.0001*** (0.00001)
Maternity Hospitals	0.030*** (0.001)	−0.005*** (0.0004)
Marital Status	0.373*** (0.003)	−0.231*** (0.001)
Insured Population	0.675*** (0.018)	−0.069*** (0.010)
Mother's education College	0.543*** (0.009)	−0.318*** (0.005)
Mother's Education Primary	0.070*** (0.008)	−0.061*** (0.005)
Mother's education High School	0.427*** (0.008)	−0.241*** (0.004)
Mother's education Ignored	0.321*** (0.014)	−0.196*** (0.008)
Mother's education None	−0.309*** (0.020)	0.157*** (0.011)
Previous Gestations	−0.207*** (0.001)	0.106*** (0.0005)
C-Section	0.421*** (0.003)	−0.110*** (0.001)
Mother's Age	0.049*** (0.0002)	−0.022*** (0.0001)
Sex Ignored	−0.574*** (0.085)	−0.095** (0.047)
Sex Male	−0.049*** (0.002)	0.016*** (0.001)
Race White	0.076*** (0.017)	−0.078*** (0.009)
Race Ignored	0.039** (0.019)	−0.044*** (0.010)
Race Indigenous	−0.537*** (0.024)	0.151*** (0.013)
Race Brown	0.028 (0.017)	−0.031*** (0.009)
Race Black	−0.048*** (0.018)	0.054*** (0.010)
Induced	0.297*** (0.003)	−0.052*** (0.002)
Fetus Presentation Pelvic	−0.338*** (0.006)	−0.071*** (0.003)
Fetus Presentation Transverse	−0.312*** (0.024)	0.037*** (0.013)
Fetus Presentation Ignored	−0.096*** (0.015)	0.058*** (0.008)
Assisted Birth Nurse	0.037*** (0.005)	0.006* (0.003)
Assisted Birth Midwive	−0.091*** (0.024)	0.044*** (0.013)
Assisted Birth Others	−0.279*** (0.023)	0.095*** (0.013)
Assisted Birth Ignored	−0.136** (0.061)	0.146*** (0.033)
Birth Place Non Hospital	−0.037 (0.028)	0.042*** (0.016)
Observations	5,042,747	5,042,747
R^2	0.046	0.035
Adjusted R^2	0.045	0.034
F Statistic (df = 28; 5037875)	8,584.424***	6,576.190***
Note:		*p<0.1; **p<0.05; ***p<0.01

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes:The table shows the first stage regression results for the 2SLS models using prenatal care visits and month of start as endogenous variables. Estimates 1 and 2 represent the results of the 2SLS models, respectively, for a sample of 1.87 million of births between 2015 and 2017. All estimates are being controlled for hospital and municipality fixed effects. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

Table 12 shows the results for the model estimation with the number of visits and month of start as the prenatal care variable and with time, municipality, and hospital fixed effects. Prenatal care visits have a mean effect of 35.4 grams for the OLS model and 70.3 grams for the instrumental variable model. The month of prenatal care start has a -5.6 statistically significant effect for the OLS model, and a -75.2 mean effect for the instrumental variable estimation.

TABLE 12 – Baseline 2SLS FE Models

	Dependent Variable			
	birth weight			
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)
Number of Visits	35.446*** (0.097)	70.257*** (1.454)	–	–
Month of Start	–	–	−5.627*** (0.179)	−75.229*** (5.092)
Time Fixed Effects	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes
Hospital Fixed Effects	Yes	Yes	Yes	Yes
R^2	0.057	0.051	0.032	0.018
Observations	5,042,747	5,042,747	5,042,747	5,042,747

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the regression results for the 2SLS models using prenatal care visits and month of start as endogenous variables. Estimates 1 and 2 represent the results for prenatal visits whereas estimates 3 and 4 for month of start. The first stage F statistics are above 10 for all models. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to 1.87 million of births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

2.6.1.1 Heterogeneity

The impact of prenatal care visits on birth weight can have greater intensity for particular groups in the sample. Running regressions for specific subsamples show how the model coefficients behave for distinct socioeconomic groups. These covariates are used as controls in previous regressions, but it is insightful to see explicitly how prenatal care coefficients change for heterogeneous groups.

TABLE 13 – Heterogenous Effects - Visits

	Dependent Variable					
	birth weight					
	Lower than Average (1)	Higher than Average (2)	Universal Health System (3)	Strictly Private Institution (4)	White (5)	Black and Brown (6)
Income	24.936*** (1.949))	80.776*** (7.092)	–	–	–	–
Institutional Setting	–	–	70.003*** (1.512)	78.998*** (4.869)	–	–
Race	–	–	–	–	82.229*** (2.654)	64.620*** (1.772)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Hospital Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.051	0.056	0.052	0.044	0.051	0.049
Observations	2,565,962	2,431,901	4,185,558	812,305	1,745,317	3,051,909

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the regression results for the 2SLS models using prenatal care visits as endogenous variable. Estimates 1 and 2 represent the results for municipalities respectively below and above the income average. Estimate 3 for births in SUS and 4 for strictly private hospitals. Estimate 5 for white and 6 for black and brown mothers. The first stage F statistics are above 10 for all models. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

In table 13 and 14, the first column refers to municipalities with below than average

GDP per capita and the second column to higher than average ones. Results indicate that the impact of prenatal visits on birth weight is higher for richer municipalities (80.7) relative to poorer ones (24.9). The same pattern happens for the prenatal delay: -117.2g for richer cities and -25.9 for poorer. One possible reason is that the quality of prenatal care is higher in more affluent cities and regions. The Unique Health System (SUS) is a decentralized health system where richer cities have the more fiscal capacity to invest in health.

The third column restricts the sample to hospitals that are part of the Unique Health System (SUS), and the fourth restricts it to strictly private hospitals. The impact is higher for strictly private hospitals, -94.268g for the prenatal delay and 78.9g for visits. A possible interpretation is related to differences in the efficiency of using health inputs (GROSSMAN, 1972). Mothers who have a birth in strictly private hospitals are part of the richest income cohorts of Brazil. They have a higher capacity to invest in other health services that are not part of prenatal care but support their health during pregnancy - like nutritionists, psychologists, and other physicians. Therefore, the marginal effect of prenatal care visits is higher than mothers who gave birth in SUS hospitals.

The fifth column restricts the sample to white mothers and the sixth to black and brown mothers. Relative to white women, black and brown mothers have a lower prenatal care impact coefficient (64.6g for visits and -66.2g for the delay). A troublesome result since there is evidence that black mothers have a higher odds of having children with low birth weight (FALCÃO et al., 2020). Again, a possible explanation for the lower impact for black mothers is that this group has, on average less education, and lack of education is a factor that decreases the efficiency of health services (GROSSMAN, 1972). Also, low birth weight children in high-risk social environments have higher persistence of poor health outcomes relative to more harmonious social environments (MCGAUHEY et al., 1991).

Table 15 show the results across different mothers´ schooling levels. Column 1 shows the results for the group with less than three years of education. Column 2 refers to the group with primary education (four to seven years). Column 3 refers to the group with high school education (eight to eleven years). And column 4 refers to the group with college-level education (more than twelve years). The results indicate that the less educated (<3 years) have a higher prenatal care impact than the others (85.6g for visits and -129.5 for the delay). The primary group has 59.5g for visits and -74.5g for the delay. The high school group had 72.4g for visits and -75.0g for the delay. Finally, the college-level group has the least impact for the delay variable (-59.7g) and 74.8g.

The prenatal care impact on birth weight is relevant in all schooling levels. The basic argument is that education increases the economic and social resources of individuals and, consequently, the capacity of using health services - for instance, being illiterate is an obstacle to the proper understanding and usage of medicine (ZIMMERMAN; WOOLF, 2014).

Table 16 shows the regression results for different Brazillian regions. The highest

TABLE 14 – Heterogenous Effects - Month of Start

	Dependent Variable					
	birth weight					
	Lower than Average (1)	Higher than Average (2)	Universal Health System (3)	Strictly Private Institution (4)	White (5)	Black and Brown (6)
Income	−25.925*** (3.229))	−117.238*** (3.835)	–	–	–	–
Institutional Setting	–	–	−71.806*** (2.629)	−94.268*** (8.283)	–	–
Race	–	–	–	–	−97.581*** (4.682)	−66.188*** (3.042)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Hospital Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.031	0.012	0.019	0.013	0.018	0.015
Observations	2,565,962	2,431,901	4,185,558	812,305	1,745,317	3,051,909

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the regression results for the 2SLS models using prenatal care month of start as endogenous variable. Estimates 1 and 2 represent the results for municipalities respectively below and above the income average. Estimate 3 for births in SUS and 4 for strictly private hospitals. Estimate 5 for white and 6 for black and brown mothers. The first stage F statistics are above 10 for all models. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 15 – Heterogeneous Effects: Schooling

	Dependent Variable			
	birth weight			
	Less than 3 years (1)	Primary (2)	High School (3)	College (4)
Visits	85.593*** (9.670)	59.560*** (3.632)	72.428*** (1.760)	74.851*** (3.761)
Month of Start	-129.573*** (19.365)	-74.496*** (6.919)	-75.000*** (3.035)	-59.732*** (6.426)
Time Fixed Effects	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes
Hospital Fixed Effects	Yes	Yes	Yes	Yes
Observations	114,341	804,839	3,016,577	1,012,588

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the regression results for the 2SLS models for different schooling subsamples. Estimates 1 represent the results for the sample with less than 3 years of schooling. Estimate 2 for primary level education. 3 for high school. And estimate 4 for mothers with college level education. The first stage F statistics are above 10 for all models. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

mean effect of visits is for the South region, 96.3g and for the delay variable the highest impact is for the Southeast, -122.7g. An interpretation is that regions have slightly different health institutional frameworks, which can be a potential driver of the results.

The negative value of prenatal care impact on the northeast is a result that needs further studies to assess its validity. However, we argue that the most plausible explanation is the Zika virus epidemic in 2015 and 2016, which mainly affected the northeastern region of Brazil (SANTOS et al., 2018). Zika virus infection is a major risk factor for the development of microcephaly in newborns (ARAÚJO et al., 2018). Mothers that had the infection will likely have more prenatal care visits but the baby´s weight will be lower - a negative self-selection

TABLE 16 – Heterogeneous Effects: Regions

	Dependent Variable				
	birth weight				
	South	Southeast	Midwest	North	Northeast
	(1)	(2)	(3)	(4)	(5)
Visits	96.348***	76.715***	61.539***	80.250***	-20.208**
	(4.845)	(1.968)	(3.335)	(5.030)	(2.622)
Month of Start	−116.374***	−122.695***	−85.950***	−50.187***	29.511
	(7.489)	(3.910)	(7.691))	(8.122)	(4.334)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes	Yes
Hospital Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	663,736	2,014,107	501,721	485,551	1,332,748

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the regression results for the 2SLS models using prenatal care visits as endogenous variable. Estimate 1 represent the results for the sample born in the southern region. Estimate 2 for the southeast. Estimate 3 for the midwest. Estimate 4 for the north. Estimate 5 for the northeast. The first stage F statistics are above 10 for all models. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

bias.

2.6.1.2 Instrument Sensitivity

An important issue to consider in instrumental variable strategies is that the bias caused by invalid instruments in two-stage least square models is proportional to the extent of overidentification in the model (ANGRIST; KRUEGER, 2001). Table 17 presents an instrument sensitivity analysis: the first model uses marital status only, the second model adds the rate of the insured instrument, and the third model adds the two other instruments (distance and hospitals per capita). The idea is to measure how the prenatal care parameter changes with the addition of new instruments to an exactly identified model. If the addition of new instruments substantially alter the prenatal coefficient, that would be evidence for distortion caused by invalid instruments. If, on the other hand, the addition of instruments do not cause the parameters to change substantially, then there is evidence for the robustness of the model to bias caused by adding weak or invalid instruments.

The estimates show that all models have statistically significant results. For the models using prenatal care visits, the difference between the overidentified model parameter (70.25) and the identified model (50.32) is around 20 grams of impact per prenatal care visit. An important difference, but one that does not change the direction of the impact - that is, more prenatal visits have a positive influence over the baby´s weight. On the other hand, for the models using the month of start prenatal variable, there is little variation in using the marital status instrumental variable only relative to using all instruments. The identified model (-87.1) and the overidentified model (-75.2) have very similar coefficients, which indicates that the overidentified model does not introduce a substantial distortion in measuring the prenatal impact.

TABLE 17 – Instrumental Variable Assessment

	Dependent Variable					
	Birth weight					
	IV: Marital Status (1)	add rate of insured (2)	All Instruments (3)	IV: Marital Status (4)	add rate of insured (5)	All Instruments (6)
Visits	54.320*** (1.538)	70.777*** (1.518)	70.257*** (1.454)	–	–	–
Month of Start	–	–	–	−87.110*** (2.540)	−87.686*** (2.541)	−75.229*** (2.513)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Hospital Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.054	0.051	0.051	0.016	0.015	0.018
Observations	5,042,747	5,042,747	5,042,747	5,042,747	5,042,747	5,042,747

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the regression results for the 2SLS models using different set of instruments for the endogeneous variable. Estimate 1 represent the results for the exactly identified model using marital status as instrument. Estimate 2 adds the rate of insured per municipality instrument. Estimate 3 is the overidentified model using all instruments. Estimate 4,5 and 6 are analogous but using month of start as the endogeneous variables. The first stage F statistics are above 10 for all models. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

2.6.2 Robustness

To assess the robustness of the impact of prenatal care to different modelling choices, a propensity score matching strategy will be used to estimate the prenatal care effect on the selected outcomes of low birth weight and very low birth weight. Matching strategies are typically used in the context of observational studies to mirror experimental studies and reduce potential biases (STUART, 2010).

Specifically, the treatment and control cohorts are built as follows: women who did an adequate amount of prenatal care (more than seven prenatal visits or started prenatal before three months of pregnancy) are in the control group ($T_i = 0$, for woman i) and women who did not do the adequate amount of prenatal care are in the treatment group ($T_i = 1$). Then the PSM approach consists of estimating the probability (a propensity score) of *not* having the adequate amount of prenatal care (T) conditional on the observed covariates (x_i) : $Pr(x_i) = Pr(T = 1|x_i)$

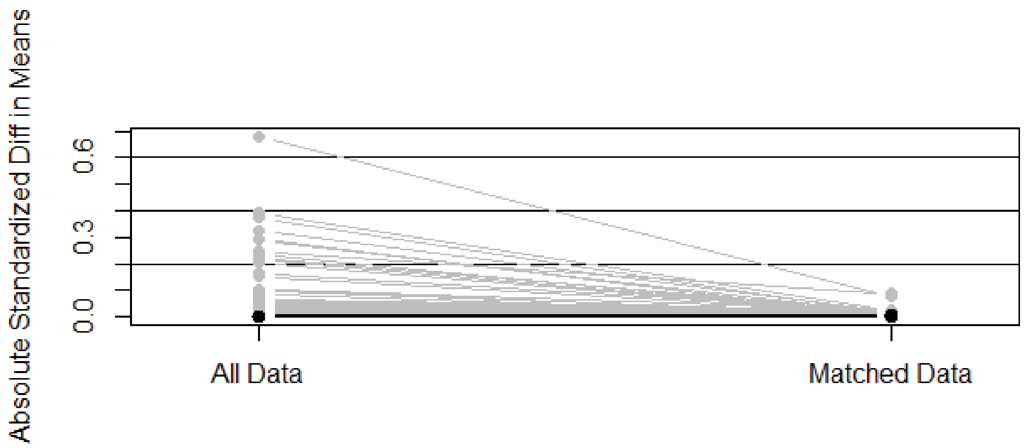
To match treated and control groups, the nearest-neighbor algorithm is used (RANDOLPH; FALBE, 2014). Each treated woman was paired with a control woman with the closest propensity score. Then, to test the quality of the match, the standard mean differences between treated and control were compared to the observed covariates included in the models. After the matching procedure, a logit model can measure how inadequate prenatal care increases the risk of low birth weight (<2500g) and very low birth weight (<1500g) in infants.

In the absence of randomization, the propensity score matching strategy is efficient in dealing with bias coming from *observable* covariates only. However, there might be *unobservable* variables that might bias the causal results between treatment and outcome variables (DIPRETE; GANGL, 2004). Therefore, to measure the sensitivity of our results to unobservable influences, we have assessed the Rosenbaum bounds of each outcome variable - a statistical test that measures the extent of the bias an unobservable variable must cause to change our results (ROSENBAUM; RUBIN, 1983).

The propensity score matching strategy´s objective is to assess the impact of inadequate prenatal care on the infant´s birth weight. To this end, two indicator variables are created to serve as outcome variables: low birth weight (if the infant is born with less than 2500g) and very low birth weight (<1500g). The matching variable, inadequate prenatal care, can be defined in two ways: less than seven prenatal visits or starting prenatal after the third month of pregnancy. Both definitions are used in the estimations.

Table 19 compares the standard mean difference between treatment and control groups after the propensity score matching procedure, using inadequate prenatal visits as treatment (matching results for prenatal delay are essentially the same, so we have omitted them). Here, the standard mean difference measures the difference in means or proportions between treated and control cohorts as percentages of standard deviations. Figure 15 summarizes the balance score for all variables. All variables are within the 0.1 threshold, which is a standard choice in the literature to guarantee common support between treatment and controls groups. That is, both groups are similar in their observable characteristics (STUART; LEE; LEACY, 2013).

FIGURE 15 – The difference in Means Before and After Matching



Source: Prepared by the authors using Unified Health System (SUS) data.

Notes: The figure shows the standardized mean difference for the treatment and control groups, before and after the propensity score matching procedure.

The observable common support between treatment and control groups allows the creation of a matched sample (of 160 thousand observations) based on the propensity scores. Then, estimating a logit model can assess how inadequate prenatal care affects the probability of low and very low birth weight in the matched sample. Results are shown in table 18. Less than seven prenatal care visits heighten the risk for newborns of having low weight [odds ratio (OR): 2.715; IR: 2.594-2.843] and very low weight [odds ratio (OR): 8.220; IR: 7.195-9.431].

Starting prenatal after three months has a lesser impact for low weight [odds ratio (OR): 1.106; IR: 1.044-1.171] and non statistically significant effect for very low weight [odds ratio (OR): 1.102; IR: 0.950-1.278].

TABLE 18 – Logit Models

	Dependent variable:			
	Low Weight		Very Low Weight	
	(1)	(2)	(3)	(4)
Treatment Visits	0.999***		2.107***	
	(0.023)		(0.069)	
Treatment Delay		0.101***		0.097
		(0.029)		(0.076)
Constant	−2.775***	−2.497***	−5.303***	−4.542***
	(0.019)	(0.021)	(0.065)	(0.055)
Observations	96,136	63,942	96,136	63,942
Log Likelihood	−30,608.870	−17,783.990	−9,466.003	−3,891.568
Akaike Inf. Crit.	61,221.740	35,571.980	18,936.010	7,787.135

Source: Prepared by the author.
Notes: The table shows the results for the Logit Model estimations in the matched dataset. Treatment visits is a dummy equal to 1 if the number of prenatal care is below seven and 0 if it is above. Treatment delay is a dummy equal to 1 if the prenatal care started after three months of pregnancy and zero otherwise.
Standard errors are in parenthesis. *p<0.1; **p<0.05; ***p<0.01

As discussed in the empirical strategy, a propensity score matching strategy can deal with bias arising from observable covariates. The Rosenbaum bounds, measuring the sensitivity of the results to unobservable influences, resulted in a gamma parameter of 1.6 to visits and 1.2 to delay. That is, even if an unobservable variable changes the odds of being affected by the treatment by a 1.6 (1.2) factor, the outcome will remain statistically significant. More intuitively, to change the model results, the unobservable covariate has to cause a change by a factor of 1.6 (1.2) in the odds of receiving inadequate prenatal care.

2.7 DISCUSSION

The FE-2SLS model results are in accordance with studies that use marital status as an instrumental variable for prenatal care and also found a positive effect on birth weight, and evidence of downward bias in OLS estimates (JEWELL; TRIUNFO, 2006). However, the mean effects of the number of prenatal visits and prenatal delay are more substantial than other studies in Brazil (WEHBY et al., 2009). The source of these differences can be the usage of distinct datasets - Unique Health System (SUS) versus alternative datasets. The SUS Live Birth System does not have detailed clinical data on mothers that, when accounted for, could reduce the impact of the prenatal care variable.

An argument is that socioeconomic characteristics - such as age, education, and income - affect only normal pregnancies, and therefore the typical variables that are used in 2SLS regression on birth weight don´t matter for *complicated* pregnancies (CONWAY; DEB, 2005). That is, it is important to distinguish normal pregnancies from pregnancies with birth defects because that can affect the prenatal care regression results (WEHBY et al., 2009). In this sense, table 29 and 35 in the appendix shows the results for regressions dividing the sample into normal births and births with a genetic anomaly: prenatal care visits and delay affect birth weight in the normal pregnancy group and the genetic anomaly group as well.

We argue that our sample size is a possible reason for the robustness of our findings to the distinction between normal and complicated pregnancies. The distinction argument is relevant when the sample size is small and outliers coming from birth defects - which causes a bimodal distribution - can bias the results. But when dealing with a dataset of millions of births, the presence of *complicated* pregnancies in the sample does not alter the results in a significant way. In the newborn context, genetics is a major factor in explaining health outcome variability, but parental choices and socioeconomic settings also matter (LEIBOWITZ, 2004).

An important point is that the fixed effects design mitigates typical concerns over external validity in the Brazillian setting. We aimed to preserve the largest sample of live births included in the microdata coming from the Brazilian vital statistical records. Another possible concern might be the interaction of setting and treatment–effects concerning different geographic zones, time periods, or institutional settings (JURAJDA, 2007). By including fixed effects from hospitals, municipalities, and time in the regressions, these possible interactions are accounted for.

The context section highlighted that Brazil has seen clear progress in reducing inequality in maternal care since the implementation of the Universal Health System in 1988 (FRANÇA et al., 2016). However, the regression results for the heterogeneous samples indicate room for substantial improvement, especially when considering the most vulnerable social groups. That is, despite substantial advances in child healthcare, major challenges remain, including the reduction of regional and socioeconomic inequalities in health (VICTORA et al., 2011). Indeed, we argue that a public policy effort to improve health endowments in fragile socioeconomic groups can have substantial welfare effects by itself, but can also contribute to the Brazilian economy by improving human capital.

Regarding the results for the schooling level regressions, there is no easy interpretation for the non-monotonic impact of prenatal visits for groups with different levels of schooling. A tentative explanation is that health efficiency gains have decreasing returns. From the group with primary education to the high school group, prenatal care increases because schooling is associated with individuals using health services more efficiently. The group with a college education has a lesser impact for prenatal because individuals already have exhausted the efficiency gains in high school - for instance, individuals do not need a college education to understand the proper usage of the medicine. However, this cannot explain why the least educated group (<3 years) has the greatest impact for visits. This mixed evidence shows that an important path for future research is to explore the interaction of education and health in

Brazil.

The robustness exercise, using the propensity score matching strategy, reinforces that inadequate prenatal care has an adverse effect on the birth weight of newborns in Brazil, (COUTINHO et al., 2009) (GONZAGA et al., 2016). Furthermore, as the matching strategy can control for observable characteristics between groups, the results are more robust to bias coming from treatment selection on observable variables, which was a limitation of studies that found mixed results for the impact of prenatal care in Brazil (VELOSO et al., 2014)(GOLDANI et al., 2004).

Finally, it is worth pointing out the limitations of the regression results. First, there is no proper measure of prenatal care quality, which can be an important source of variability in the results. Second, there is no information regarding smoking and drinking habits, which can also explain variability in birth weights. However, these limitations do not alter the general message of our findings that prenatal care is an important factor in improving the health results of newborns in Brazil.

2.8 FINAL REMARKS AND PUBLIC POLICY IMPLICATIONS

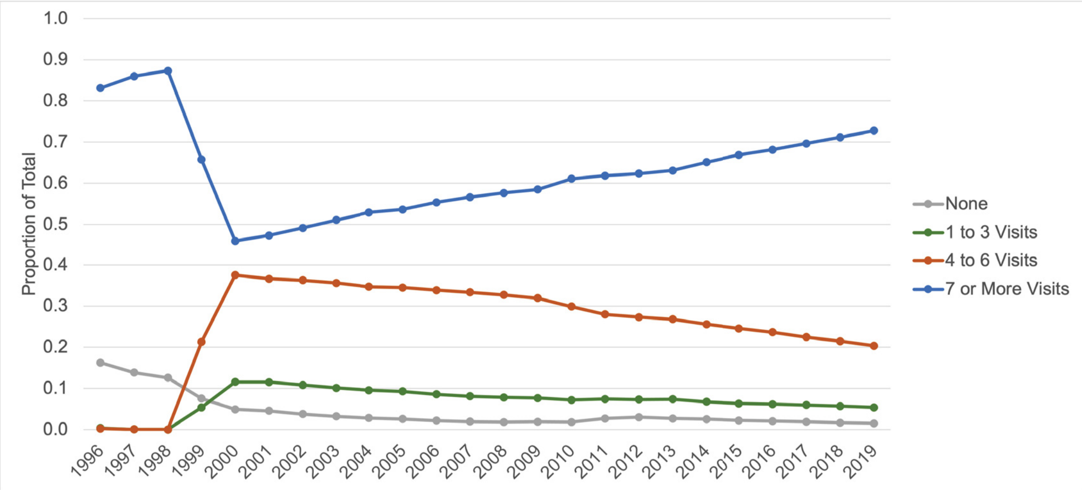
Adverse health conditions at birth can have long-lasting consequences for the newborn´s life. Infants with low birth weight have higher odds of mortality and impairments to their full cognitive development. In developing countries, health problems in infancy can be an obstacle that helps perpetuate states of poverty in individuals of fragile socioeconomic groups.

The paper´s contribute to the objective of identifying the impact of prenatal care on newborn weight in the Brazillian context - a continental middle-income country with substantial regional heterogeneity within its borders. In particular, the instrumental variable results show that an increase in prenatal care visits positively affects the baby´s weight. In contrast, a delay in the prenatal start has a negative effect. Results are robust to weak instruments problem and heterogeneity. In turn, the propensity score matching results indicate that inadequate prenatal care increases the odds of a newborn having low weight and very low birth weight. This result is robust to self-selection in the observable characteristics and up to an important degree in unobservable variables as well.

The public policy implication is that increasing prenatal care among women - especially in poorer regions and municipalities - will cause an increase in average birth weight, which can improve newborns' health conditions. That is important because, as discussed, improving newborn health conditions can have long-term impacts on income, cognitive capacity, and health in general. Providing quality prenatal care can be a key part of public health strategies that improve infant´s life in middle-income countries.

2.9 APPENDIX 2:

FIGURE 16 – Prenatal Care Time Series



Source: Prepared by the authors using the Unified Health System (SUS) data.

Notes: The figure shows the evolution of prenatal care visits proportions in Brazil, between 1996 and 2019.

TABLE 19 – Covariate Balance Matched Sample

	Matched Data		
Variable	**Means Treated**	**Means Control**	**Std. Mean Diff.**
Distance	0.287	0.275	0.071
Mother's Education College	0.171	0.167	0.011
Mother's Education Primary	0.184	0.186	-0.005
Mother's Education High School	0.615	0.618	-0.005
Mother's Education Ignored	0.006	0.006	0.000
Mother's Education None	0.004	0.004	0.002
Type of birth	0.461	0.457	0.008
Mother Age	26.058	25.987	0.010
Father Age	29.810	29.711	0.012
Birth Location	0.010	0.010	0.006
Gestational Weeks	37.750	38.010	-0.087
Previous Gestations	1.330	1.318	0.008
Vaginal Births	0.792	0.758	0.025
Cesarean Births	0.390	0.395	-0.007
Parity Double	0.026	0.024	0.013
Parity Ignored	0.000	0.000	0.000
Parity Triple	0.001	0.001	0.007
Race White	0.403	0.401	0.005
Race Ignored	0.017	0.017	0.001
Race Indigenous	0.010	0.009	0.008
Race Brown	0.498	0.502	-0.006
Race Black	0.066	0.067	-0.004
Marital Status Ignored	0.006	0.005	0.014
Marital Status Divorced	0.012	0.014	-0.020
Marital Status Single	0.444	0.445	-0.004
Marital Status Common-law Marriage	0.239	0.240	-0.002
Marital Status Widower	0.001	0.002	-0.003
Sex Ignored	0.000	0.000	0.012
Sex male	0.518	0.523	-0.009
Fetus Presentation	1.116	1.113	0.004
Induced Labor	1.942	1.942	0.000
C section before start	2.312	2.313	-0.001
Assisted labor	1.096	1.098	-0.003

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows means for a set of maternal and pregnancy characteristics by time of day and the standard mean difference for two groups: women who had birth at the Brazillian commercial time (between 08:00 and 12:00 and 14:00 to 18:00) versus women who had birth in the rest of the day. The sample is restricted to single births, without genetic anomalies and performed in public hospitals.*p<0.1; **p<0.05; ***p<0.01

TABLE 20 – Baseline Fixed Effects Model - Visits

	Dependent variable:	
	Birth Weight	
	(1)	(2)
Prenatal Visits	35.446*** (0.097)	70.257*** (1.454)
Previous Gestations	26.919*** (0.191)	34.379*** (0.366)
Mother's Age	0.435*** (0.044)	−1.476*** (0.091)
Sex Ignored	−581.451*** (18.510)	−561.081*** (18.763)
Sex Male	108.035*** (0.472)	109.746*** (0.483)
Race White	7.211* (3.771)	4.230 (3.821)
Race Ignored	12.810*** (4.141)	11.965*** (4.193)
Race Indigenous	34.961*** (5.326)	53.748*** (5.449)
Race Brown	11.903*** (3.766)	11.103*** (3.813)
Race Black	−3.364 (3.882)	−1.358 (3.932)
Mother's education College	35.987*** (1.883)	13.926*** (2.117)
Mother's Education Primary	28.301*** (1.809)	25.396*** (1.836)
Mother's education High School	48.323*** (1.767)	31.932*** (1.915)
Mother's education Ignored	31.921*** (3.163)	19.649*** (3.243)
Mother's education None	−37.083*** (4.470)	−25.639*** (4.551)
Induced	73.551*** (0.699)	63.290*** (0.827)
C-Section	86.708*** (0.586)	71.716*** (0.862)
Fetus Presentation Pelvic	−290.719*** (1.279)	−278.866*** (1.386)
Fetus Presentation Transverse	−313.995*** (5.279)	−302.864*** (5.366)
Fetus Presentation Ignored	−80.407*** (3.228)	−76.990*** (3.272)
Assisted Birth Nurse	37.919*** (1.201)	36.621*** (1.218)
Assisted Birth Midwive	−0.210 (5.289)	2.875 (5.357)
Assisted Birth Others	−7.879 (5.060)	1.887 (5.140)
Assisted Birth Ignored	−31.538** (13.265)	−26.747** (13.434)
Birth Place Non Hospital		−9.021 (6.227)
−7.759 (6.306)		
Observations	5,042,747	5,042,747
R^2	0.057	0.051
Adjusted R^2	0.056	0.050
F Statistic	12,099.430*** (df = 25; 5037878)	167,972.100***
Note:		*p<0.1; **p<0.05; ***p<0.01

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the regression results for the 2SLS models using prenatal care visits as endogenous variables. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to 1.87 million of births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 21 – Baseline Fixed Effects Model - Month of Start

	Dependent variable:	
	Birth Weight	
	(1)	(2)
Month of Start	−5.627*** (0.179)	−75.229*** (2.513)
Previous Gestations	19.946*** (0.193)	27.645*** (0.339)
C-Section	101.313*** (0.592)	93.145*** (0.669)
Mother's Age	2.238*** (0.045)	0.475*** (0.078)
Sex Ignored	−602.735*** (18.750)	−609.450*** (19.032)
Sex Male	106.383*** (0.478)	107.504*** (0.487)
Race White	9.772** (3.820)	3.903 (3.883)
Race Ignored	13.463*** (4.194)	10.906** (4.258)
Race Indigenous	16.682*** (5.395)	27.198*** (5.488)
Race Brown	12.553*** (3.815)	10.507*** (3.872)
Race Black	−5.071 (3.933)	−0.920 (3.994)
Mother's education College	56.363*** (1.908)	30.546*** (2.148)
Mother's Education Primary	30.874*** (1.833)	26.114*** (1.868)
Mother's education High School	63.513*** (1.790)	44.978*** (1.935)
Mother's education Ignored	43.210*** (3.204)	28.283*** (3.296)
Mother's education None	−47.793*** (4.528)	−36.145*** (4.615)
Induced	83.712*** (0.707)	80.155*** (0.729)
Fetus Presentation Pelvic	−303.185*** (1.295)	−308.091*** (1.327)
Fetus Presentation Transverse	−325.102*** (5.347)	−322.288*** (5.428)
Fetus Presentation Ignored	−83.555*** (3.270)	−79.460*** (3.322)
Assisted Birth Nurse	39.272*** (1.217)	39.647*** (1.235)
Assisted Birth Midwive	−3.109 (5.358)	−0.114 (5.439)
Assisted Birth Others	−17.288*** (5.126)	−10.670** (5.208)
Assisted Birth Ignored	−35.580*** (13.438)	−25.228* (13.643)
Birth Place Non Hospital	−7.102 (6.403)	−10.066 (6.308)
Observations	5,042,747	5,042,747
R^2	0.032	0.018
Adjusted R^2	0.031	0.017
F Statistic	6,660.411*** (df = 25; 5037878)	161,583.200***
Note:		*p<0.1; **p<0.05; ***p<0.01

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the regression results for the 2SLS models using prenatal care month of start as endogenous variables. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to 1.87 million of births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 22 – Instrument Assessment - Visits

	Dependent Variable: Birth Weight		
	IV: Marital Status	add Rate of Insured	All instruments
	(1)	(2)	(3)
Prenatal Visits	54.320*** (1.538)	70.777*** (1.518)	70.257*** (1.454)
Mother's education College	24.026*** (2.126)	13.597*** (2.135)	13.926*** (2.117)
Mother's Education Primary	26.726*** (1.820)	25.353*** (1.837)	25.396*** (1.836)
Mother's education High School	39.436*** (1.915)	31.687*** (1.927)	31.932*** (1.915)
Mother's education Ignored	25.267*** (3.220)	19.466*** (3.248)	19.649*** (3.243)
Mother's education None	−30.878*** (4.515)	−25.468*** (4.555)	−25.639*** (4.551)
Previous Gestations	30.964*** (0.381)	34.491*** (0.378)	34.379*** (0.366)
C-Section	78.580*** (0.885)	71.492*** (0.882)	71.716*** (0.862)
Mother's Age	−0.601*** (0.095)	−1.505*** (0.094)	−1.476*** (0.091)
Sex Ignored	−570.407*** (18.601)	−560.777*** (18.771)	−561.081*** (18.763)
Sex Male	108.963*** (0.480)	109.772*** (0.484)	109.746*** (0.483)
Race White	5.594 (3.787)	4.185 (3.822)	4.230 (3.821)
Race Ignored	12.352*** (4.156)	11.952*** (4.194)	11.965*** (4.193)
Race Indigenous	45.147*** (5.409)	54.028*** (5.456)	53.748*** (5.449)
Race Brown	11.469*** (3.780)	11.091*** (3.815)	11.103*** (3.813)
Race Black	−2.276 (3.898)	−1.328 (3.934)	−1.358 (3.932)
Induced	67.987*** (0.835)	63.136*** (0.837)	63.290*** (0.827)
Fetus Presentation Pelvic	−284.292*** (1.386)	−278.689*** (1.395)	−278.866*** (1.386)
Fetus Presentation Transverse	−307.960*** (5.321)	−302.697*** (5.369)	−302.864*** (5.366)
Fetus Presentation Ignored	−78.554*** (3.244)	−76.939*** (3.274)	−76.990*** (3.272)
Assisted Birth Nurse	37.216*** (1.207)	36.602*** (1.218)	36.621*** (1.218)
Assisted Birth Midwive	1.462 (5.311)	2.921 (5.359)	2.875 (5.357)
Assisted Birth Others	−2.584 (5.098)	2.033 (5.144)	1.887 (5.140)
Assisted Birth Ignored	−28.941** (13.317)	−26.676** (13.439)	−26.747** (13.434)
Birth Place Non Hospital		−8.337 (6.250)	−7.740 (6.308)
−7.759 (6.306)			
Observations	5,042,747	5,042,747	5,042,747
R^2	0.054	0.051	0.051
Adjusted R^2	0.053	0.050	0.050
F Statistic	169,833.000***	167,689.500***	167,972.100***

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the instrument assessment results for the 2SLS models using prenatal care visits as endogenous variables. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to 1.87 million of births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 23 – Instrument Assessment - Month of Start

	Dependent Variable: Newborn Weight		
	IV: Marital Status	add Rate of Insured	All instruments
	(1)	(2)	(3)
Month of Start	−87.110*** (2.540)	−87.686*** (2.541)	−75.229*** (2.513)
Mother's education College	26.139*** (2.162)	25.925*** (2.162)	30.546*** (2.148)
Mother's Education Primary	25.301*** (1.878)	25.262*** (1.879)	26.114*** (1.868)
Mother's education High School	41.814*** (1.947)	41.660*** (1.948)	44.978*** (1.935)
Mother's education Ignored	25.735*** (3.314)	25.611*** (3.315)	28.283*** (3.296)
Mother's education None	−34.156*** (4.640)	−34.060*** (4.641)	−36.145*** (4.615)
Previous Gestations	28.959*** (0.343)	29.023*** (0.343)	27.645*** (0.339)
C-Section	91.750*** (0.674)	91.683*** (0.674)	93.145*** (0.669)
Mother's Age	0.174** (0.079)	0.159** (0.079)	0.475*** (0.078)
Sex Ignored	−610.596*** (19.134)	−610.651*** (19.140)	−609.450*** (19.032)
Sex Male	107.695*** (0.489)	107.705*** (0.490)	107.504*** (0.487)
Race White	2.902 (3.904)	2.853 (3.905)	3.903 (3.883)
Race Ignored	10.470** (4.281)	10.448** (4.282)	10.906** (4.258)
Race Indigenous	28.993*** (5.518)	29.080*** (5.519)	27.198*** (5.488)
Race Brown	10.158*** (3.893)	10.141*** (3.894)	10.507*** (3.872)
Race Black	−0.211 (4.016)	−0.177 (4.017)	−0.920 (3.994)
Induced	79.548*** (0.733)	79.519*** (0.733)	80.155*** (0.729)
Fetus Presentation Pelvic	−308.928*** (1.334)	−308.969*** (1.334)	−308.091*** (1.327)
Fetus Presentation Transverse	−321.807*** (5.457)	−321.784*** (5.459)	−322.288*** (5.428)
Fetus Presentation Ignored	−78.761*** (3.340)	−78.727*** (3.341)	−79.460*** (3.322)
Assisted Birth Nurse	39.711*** (1.242)	39.714*** (1.242)	39.647*** (1.235)
Assisted Birth Midwive	0.398 (5.468)	0.422 (5.470)	−0.114 (5.439)
Assisted Birth Others	−9.540* (5.236)	−9.485* (5.238)	−10.670** (5.208)
Assisted Birth Ignored	−23.461* (13.717)	−23.375* (13.721)	−25.228* (13.643)
Birth Place Non Hospital		−6.596 (6.437)	−6.571 (6.439)
−7.102 (6.403)			
Observations	5,042,747	5,042,747	5,042,747
R^2	0.016	0.015	0.018
Adjusted R^2	0.015	0.015	0.017
F Statistic	160,143.000***	160,068.600***	161,583.200***

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the instrument assessment results for the 2SLS models using prenatal care delay as endogenous variables. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to 1.87 million of births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 24 – Heterogenous effects: Municipalities´ Income

	Dependent Variable: Birth Weight	
	Lower than Median	Higher than Median
	(1)	(2)
Prenatal Visits	24.936*** (1.949)	80.776*** (1.848)
Previous Gestations	24.275*** (0.471)	37.751*** (0.512)
C-Section	106.751*** (1.205)	51.316*** (1.141)
Mother's Age	1.696*** (0.121)	−2.828*** (0.123)
Sex Ignored	−529.332*** (25.024)	−616.322*** (28.339)
Sex Male	106.365*** (0.662)	111.718*** (0.705)
Race White	−11.060* (6.192)	12.742*** (4.876)
Race Ignored	0.348 (6.561)	13.402** (5.586)
Race Indigenous	−3.300 (7.614)	123.108*** (9.075)
Race Brown	−3.916 (6.155)	18.478*** (4.884)
Race Black	−9.855 (6.320)	−1.351 (5.052)
Mother's education College	32.375*** (2.653)	27.085*** (3.744)
Mother's Education Primary	31.452*** (2.128)	29.412*** (3.585)
Mother's education High School	50.230*** (2.290)	39.880*** (3.576)
Mother's education Ignored	32.662*** (3.759)	29.338*** (6.305)
Mother's education None	−45.504*** (5.000)	−3.438 (10.537)
Induced	72.353*** (1.155)	59.697*** (1.140)
Fetus Presentation Pelvic	−283.803*** (1.955)	−282.349*** (1.933)
Fetus Presentation Transverse	−247.116*** (7.264)	−370.010*** (7.926)
Fetus Presentation Ignored	−73.955*** (4.684)	−81.100*** (4.569)
Assisted Birth Nurse	20.035*** (1.684)	50.646*** (1.789)
Assisted Birth Midwive	2.068 (5.405)	9.420 (24.979)
Assisted Birth Others	−3.302 (5.294)	−52.230*** (17.627)
Assisted Birth Ignored	−23.223 (17.980)	−43.993** (20.100)
Birth Place Non Hospital		−15.543* (8.010)
−25.764* (13.349)		
Observations	2,565,962	2,431,901
R^2	0.051	0.056
Adjusted R^2	0.050	0.056
F Statistic	90,208.740***	79,907.600***

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the results for the 2SLS models using prenatal care delay as endogenous variables. Column 1 refers to the sample of births in lower than average municipalities and column 2 to the sample of higher than average ones. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 25 – Heterogenous effects: Institutional Setting

	Dependent Variable: Birth Weight	
	Universal Health System	Strictly Private
	(1)	(2)
Prenatal Visits	70.003*** (1.512)	78.988*** (4.869)
Previous Gestations	34.360*** (0.400)	33.549*** (0.834)
C-Section	71.299*** (0.919)	67.987*** (2.263)
Mother's Age	$-$1.272*** (0.102)	$-$2.872*** (0.178)
Sex Ignored	$-$583.108*** (19.917)	$-$317.831*** (60.457)
Sex Male	108.839*** (0.534)	114.632*** (1.170)
Race White	$-$6.140 (4.643)	26.278*** (6.642)
Race Ignored	0.619 (4.990)	35.957*** (8.228)
Race Indigenous	42.073*** (6.094)	80.385*** (21.755)
Race Brown	$-$0.928 (4.621)	41.277*** (6.727)
Race Black	$-$13.561*** (4.736)	30.326*** (7.236)
Mother's education College	13.291*** (2.194)	2.619 (12.047)
Mother's Education Primary	26.447*** (1.878)	25.627** (12.073)
Mother's education High School	33.524*** (1.964)	15.692 (11.741)
Mother's education Ignored	18.945*** (3.361)	36.432** (15.745)
Mother's education None	$-$26.951*** (4.632)	60.293* (36.412)
Induced	66.187*** (0.890)	35.170*** (2.247)
Fetus Presentation Pelvic	$-$283.688*** (1.538)	$-$256.163*** (3.357)
Fetus Presentation Transverse	$-$297.266*** (5.952)	$-$322.201*** (12.608)
Fetus Presentation Ignored	$-$81.820*** (3.674)	$-$51.915*** (7.258)
Assisted Birth Nurse	36.082*** (1.260)	14.996** (7.288)
Assisted Birth Midwive	2.415 (5.425)	$-$7.933 (59.239)
Assisted Birth Others	1.512 (5.223)	$-$11.535 (40.789)
Assisted Birth Ignored	$-$28.226* (14.630)	$-$6.249 (34.714)
Birth Place Non Hospital		$-$16.118** (7.980)
$-$22.867 (13.960)		
Observations	4,185,558	812,305
R^2	0.052	0.044
Adjusted R^2	0.051	0.043
F Statistic	143,560.800***	23,593.040***

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the results for the 2SLS models using prenatal care visit as endogenous variables. Column 1 refers to the sample of births in SUS health units and column 2 to strictly private ones. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 26 – Heterogenous effects: Ethnicity

	Dependent Variable: Birth Weight	
	White	Black and Brown
	(1)	(2)
Prenatal Visits	82.229*** (2.654)	64.260*** (1.772)
Previous Gestations	41.293*** (0.667)	31.278*** (0.456)
C-Section	59.063*** (1.457)	78.577*** (1.094)
Mother's Age	−3.301*** (0.147)	−0.425*** (0.119)
Sex Ignored	−638.110*** (36.109)	−561.684*** (23.772)
Sex Male	114.703*** (0.822)	106.969*** (0.622)
Mother's education College	21.572*** (4.826)	14.911*** (2.563)
Mother's Education Primary	21.892*** (4.666)	28.726*** (2.095)
Mother's education High School	33.837*** (4.650)	34.021*** (2.219)
Mother's education Ignored	26.853*** (7.933)	19.968*** (3.860)
Mother's education None	−26.437* (13.680)	−27.662*** (5.204)
Induced	62.841*** (1.452)	62.316*** (1.038)
Fetus Presentation Pelvic	−272.562*** (2.283)	−279.158*** (1.825)
Fetus Presentation Transverse	−358.790*** (8.712)	−268.362*** (7.114)
Fetus Presentation Ignored	−72.417*** (5.930)	−73.556*** (4.130)
Assisted Birth Nurse	38.493*** (2.546)	36.463*** (1.454)
Assisted Birth Midwive	−8.966 (17.245)	4.406 (5.919)
Assisted Birth Others	4.388 (14.769)	0.973 (5.719)
Assisted Birth Ignored	−49.699* (27.695)	−18.540 (16.431)
Birth Place Non Hospital		−33.582*** (13.023)
−11.833 (8.566)		
Observations	1,745,317	3,051,909
R^2	0.053	0.050
Adjusted R^2	0.051	0.049
F Statistic	58,659.460***	101,597.000***

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the results for the 2SLS models using prenatal care visit as endogenous variables. Column 1 refers to the sample of white mothers and column 2 to black and brown ones. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level.
*p<0.1; **p<0.05; ***p<0.01

TABLE 27 – Heterogenous effects: Mother's Education

	Dependent Variable: Birth Weight			
	Primary	Secondary	High School	College
	(1)	(2)	(3)	(4)
Prenatal Visits	85.593*** (9.670)	59.560*** (3.632)	72.428*** (1.760)	74.851*** (3.761)
Previous Gestations	23.957*** (1.836)	27.776*** (0.926)	36.315*** (0.493)	40.112*** (0.744)
C-Section	35.540*** (6.748)	65.743*** (2.349)	74.607*** (1.079)	70.282*** (1.790)
Mother's Age	−2.327*** (0.589)	0.274 (0.273)	−1.114*** (0.115)	−4.355*** (0.151)
Sex Ignored	−762.651*** (117.488)	−474.845*** (43.578)	−582.308*** (23.747)	−560.942*** (52.119)
Sex Male	103.022*** (3.430)	107.117*** (1.222)	109.348*** (0.625)	113.944*** (1.066)
Race White	−4.826 (33.857)	12.778 (11.958)	−10.379* (5.307)	25.263*** (6.390)
Race Ignored	8.874 (35.145)	14.777 (12.572)	−2.344 (5.782)	34.293*** (7.629)
Race Indigenous	10.988 (36.348)	34.589** (14.009)	48.635*** (7.653)	108.133*** (17.027)
Race Brown	−6.814 (33.534)	15.001 (11.874)	−5.148 (5.282)	40.690*** (6.456)
Race Black	−22.929 (34.150)	−0.088 (12.094)	−16.745*** (5.417)	28.287*** (6.982)
Induced	75.878*** (6.169)	70.408*** (2.087)	63.303*** (1.028)	52.419*** (1.965)
Fetus Presentation Pelvic	−252.992*** (9.986)	−279.336*** (3.663)	−285.752*** (1.843)	−266.297*** (2.783)
Fetus Presentation Transverse	−146.800*** (31.867)	−266.212*** (13.241)	−295.686*** (7.239)	−372.853*** (11.180)
Fetus Presentation Ignored	−37.308 (23.456)	−89.859*** (8.248)	−74.617*** (4.420)	−73.355*** (6.761)
Assisted Birth Nurse	21.222*** (7.590)	34.690*** (2.774)	38.640*** (1.507)	29.722*** (4.263)
Assisted Birth Midwive	−12.118 (20.152)	−0.195 (9.942)	3.331 (7.303)	0.014 (25.541)
Assisted Birth Others	−24.694 (21.425)	−1.117 (9.767)	3.341 (6.858)	−27.794 (20.656)
Assisted Birth Ignored	77.352 (74.453)	−47.256 (29.980)	−24.910 (17.672)	−1.926 (35.112)
Birth Place Non Hospital		−0.876 (44.904)	−36.477** (16.611)	−13.088 (9.137)
−24.132 (15.870)				
Observations	114,341	804,839	3,016,577	1,012,588
R^2	0.042	0.054	0.052	0.047
Adjusted R^2	0.016	0.049	0.050	0.044
F Statistic	2,935.308***	27,868.100***	102,680.600***	32,648.160***

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the results for the 2SLS models using prenatal care visit as endogenous variables. Column 1 refers to the sample of women with less than three years of education. In column 2, four to seven years of schooling. In column 3, eight to eleven years of schooling. In column 4, more than twelve years of schooling. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 28 – Heterogenous effects: Region of Birth

	Dependent Variable: Birth Weight				
	South	Southeast	MidWest	North	Northeast
	(1)	(2)	(3)	(4)	(5)
Prenatal Visits	96.348*** (4.845)	76.715*** (1.968)	61.539*** (3.335)	80.250*** (5.030)	−20.208*** (2.622)
Previous Gestations	50.738*** (1.376)	37.357*** (0.589)	23.563*** (0.765)	44.903*** (1.392)	13.974*** (0.585)
C-Section	49.398*** (2.558)	57.959*** (1.266)	56.479*** (2.418)	60.969*** (2.963)	140.458*** (1.622)
Mother's Age	−4.932*** (0.297)	−2.714*** (0.132)	−0.359 (0.221)	−2.036*** (0.414)	4.621*** (0.152)
Sex Ignored	−616.473*** (72.252)	−580.260*** (29.912)	−487.295*** (58.951)	−507.328*** (79.312)	−595.885*** (31.602)
Sex Male	115.592*** (1.379)	110.899*** (0.765)	110.010*** (1.471)	105.661*** (1.562)	103.328*** (0.963)
Race White	20.755 (14.115)	12.215** (5.370)	−7.986 (9.727)	−24.934 (16.054)	−12.447 (8.644)
Race Ignored	19.473 (16.263)	16.388** (6.405)	9.863 (10.238)	−2.416 (17.139)	−1.297 (9.018)
Race Indigenous	137.888*** (21.796)	187.036*** (12.073)	60.646*** (13.334)	−3.207 (17.161)	−32.100*** (12.096)
Race Brown	23.320 (14.207)	16.726*** (5.379)	10.323 (9.650)	−7.789 (15.883)	−4.105 (8.545)
Race Black	21.723 (14.505)	−4.911 (5.522)	9.103 (10.356)	−4.541 (16.751)	−3.644 (8.787)
Mother's education College	49.002*** (6.898)	32.944*** (4.019)	−20.322*** (7.284)	−12.476* (6.923)	55.080*** (3.509)
Mother's Education Primary	41.619*** (6.482)	30.341*** (3.845)	9.664 (6.633)	25.313*** (4.758)	37.201*** (2.823)
Mother's education High School	56.800*** (6.647)	43.951*** (3.831)	8.919 (6.673)	15.761*** (5.576)	69.813*** (3.019)
Mother's education Ignored	42.450*** (15.394)	32.345*** (6.427)	1.372 (10.975)	3.892 (9.049)	39.987*** (4.908)
Mother's education None	−52.454*** (18.428)	−23.143** (11.559)	7.134 (16.637)	−24.816** (10.110)	−66.405*** (6.711)
Induced	76.036*** (2.605)	54.494*** (1.225)	55.263*** (2.400)	50.310*** (2.887)	77.854*** (1.659)
Fetus Presentation Pelvic	−268.017*** (3.803)	−295.335*** (2.205)	−219.125*** (4.129)	−198.636*** (4.611)	−326.395*** (2.751)
Fetus Presentation Transverse	−343.018*** (12.788)	−363.876*** (8.916)	−290.758*** (19.379)	−156.754*** (16.538)	−258.487*** (10.671)
Fetus Presentation Ignored	−64.586*** (10.289)	−95.527*** (5.985)	−78.332*** (6.872)	−39.804*** (7.503)	−83.713*** (7.850)
Assisted Birth Nurse	26.791*** (5.686)	58.039*** (1.844)	45.557*** (6.765)	11.350*** (3.174)	23.476*** (2.347)
Assisted Birth Midwive	−45.189 (38.729)	6.038 (27.118)	−3.396 (50.742)	13.312 (8.970)	−6.094 (7.406)
Assisted Birth Others	−41.977 (33.566)	−38.604** (16.012)	−29.519 (30.120)	4.811 (8.390)	−15.886** (7.900)
Assisted Birth Ignored	33.027 (55.330)	−55.780** (27.146)	−40.123 (26.635)	−41.731 (36.081)	−27.766 (24.787)
Birth Place Non Hospital		−33.417 (22.825)	−30.605** (13.188)	9.699 (25.331)	0.547 (20.266)
−18.055* (10.828)					
Observations	663,736	2,014,107	501,721	485,551	1,332,748
R^2	0.053	0.059	0.037	0.032	0.014
Adjusted R^2	0.052	0.058	0.036	0.030	0.013
F Statistic	23,664.390***	66,993.000***	13,916.930***	15,192.860***	48,809.050***

Source: Prepared by the authors using Unique Health System (SUS) data.

Notes: The table shows the results for the 2SLS models using prenatal care visits as endogenous variable. Estimate 1 represent the results for the sample born in the southern region. Estimate 2 for the southeast. Estimate 3 for the midwest. Estimate 4 for the north. Estimate 5 for the northeast. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 29 – Heterogenous effects: Genetic Anomaly

	Dependent Variable: Birth Weight	
	Normal	Genetic Anomaly
	(1)	(2)
Prenatal Visits	69.559*** (1.496)	63.039*** (18.036)
Previous Gestations	35.981*** (0.390)	21.877*** (3.520)
C-Section	72.171*** (0.876)	102.494*** (13.124)
Mother's Age	−1.595*** (0.095)	−3.026*** (0.983)
Sex Ignored		−438.467*** (30.859)
Sex Male	110.066*** (0.490)	132.434*** (6.837)
Race White	4.978 (3.896)	−7.107 (43.256)
Race Ignored	12.641*** (4.296)	18.999 (50.459)
Race Indigenous	54.117*** (5.520)	123.148* (72.965)
Race Brown	11.594*** (3.889)	18.724 (43.322)
Race Black	−1.300 (4.009)	17.314 (44.896)
Mother's education College	18.302*** (2.142)	−37.747 (29.925)
Mother's Education Primary	27.352*** (1.864)	−28.247 (25.872)
Mother's education High School	35.287*** (1.941)	−15.766 (27.464)
Mother's education Ignored	23.144*** (3.338)	−55.228 (52.851)
Mother's education None	−25.364*** (4.609)	−110.337* (62.025)
Induced	62.530*** (0.841)	118.644*** (11.717)
Fetus Presentation Pelvic	−274.688*** (1.412)	−343.028*** (15.032)
Fetus Presentation Transverse	−296.790*** (5.440)	−530.478*** (57.996)
Fetus Presentation Ignored	−82.010*** (3.566)	−205.536*** (43.434)
Assisted Birth Nurse	35.233*** (1.240)	82.177*** (17.730)
Assisted Birth Midwive	−0.091 (5.391)	114.665 (112.177)
Assisted Birth Others	0.478 (5.179)	94.762 (88.567)
Assisted Birth Ignored	−17.474 (14.381)	44.153 (205.184)
Birth Place Non Hospital		−10.272 (7.088)
−104.485* (56.274)		
Observations	4,859,663	44,726
R^2	0.050	0.090
Adjusted R^2	0.049	0.042
F Statistic	160,402.200***	2,277.508***

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the results for the 2SLS models using prenatal care visits as endogenous variable. Column 1 refers to the sample of normal births and column 2 births with genetic anomaly. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 30 – Heterogenous effects: Municipalities´ Income - Month of Start

	Dependent Variable: Birth Weight	
	Lower than Median	Higher than Median
	(1)	(2)
Month of Start	−25.925*** (3.229)	−117.238*** (3.835)
Previous Gestations	22.086*** (0.443)	32.244*** (0.516)
C-Section	114.609*** (0.930)	72.560*** (0.968)
Mother's Age	2.326*** (0.107)	−1.015*** (0.112)
Sex Ignored	−544.336*** (25.278)	−684.474*** (29.208)
Sex Male	105.681*** (0.666)	109.176*** (0.722)
Race White	−11.887* (6.262)	11.065** (5.039)
Race Ignored	−1.018 (6.632)	13.034** (5.764)
Race Indigenous	−14.339* (7.625)	108.203*** (9.357)
Race Brown	−5.006 (6.221)	18.733*** (5.039)
Race Black	−10.426 (6.392)	1.692 (5.214)
Mother's education College	38.894*** (2.667)	33.879*** (3.913)
Mother's Education Primary	32.092*** (2.155)	26.365*** (3.702)
Mother's education High School	55.430*** (2.283)	45.658*** (3.712)
Mother's education Ignored	36.270*** (3.797)	30.715*** (6.523)
Mother's education None	−49.162*** (5.044)	−15.839 (10.862)
Induced	77.021*** (1.076)	81.238*** (1.010)
Fetus Presentation Pelvic	−294.130*** (1.900)	−317.598*** (1.883)
Fetus Presentation Transverse	−251.960*** (7.330)	−397.642*** (8.140)
Fetus Presentation Ignored	−71.952*** (4.761)	−93.346*** (4.711)
Assisted Birth Nurse	21.200*** (1.705)	53.559*** (1.843)
Assisted Birth Midwive	0.900 (5.463)	11.761 (25.769)
Assisted Birth Others	−7.371 (5.336)	−73.737*** (18.171)
Assisted Birth Ignored	−25.931 (18.175)	−24.538 (20.743)
Birth Place Non Hospital		−15.796* (8.097)
−18.285 (13.771)		
Observations	2,565,962	2,431,901
R^2	0.031	0.012
Adjusted R^2	0.029	0.011
F Statistic	88,181.430***	74,223.940***

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the results for the 2SLS models using prenatal care delay as endogenous variables. Column 1 refers to the sample of births in lower than average municipalities and column 2 to the sample of higher than average ones. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 31 – Heterogenous effects: Institutional Setting - Month of Start

	Dependent Variable: Birth Weight	
	Universal Health System	Strictly Private
	(1)	(2)
Month of Start	−71.806*** (2.629)	−94.268*** (8.283)
Previous Gestations	26.946*** (0.374)	28.454*** (0.717)
C-Section	93.521*** (0.715)	84.683*** (1.923)
Mother's Age	0.935*** (0.086)	−2.058*** (0.169)
Sex Ignored	−631.891*** (20.181)	−365.429*** (61.313)
Sex Male	106.535*** (0.538)	112.215*** (1.173)
Race White	−4.634 (4.712)	20.808*** (6.755)
Race Ignored	1.394 (5.062)	31.541*** (8.345)
Race Indigenous	17.015*** (6.143)	56.799*** (22.006)
Race Brown	0.723 (4.687)	33.789*** (6.809)
Race Black	−10.997** (4.804)	22.406*** (7.315)
Mother's education College	28.574*** (2.251)	27.682** (12.025)
Mother's Education Primary	27.738*** (1.908)	28.442** (12.258)
Mother's education High School	47.639*** (1.981)	28.954** (11.862)
Mother's education Ignored	27.690*** (3.414)	60.481*** (15.854)
Mother's education None	−38.381*** (4.690)	56.659 (36.969)
Induced	83.949*** (0.780)	45.305*** (2.191)
Fetus Presentation Pelvic	−313.103*** (1.480)	−287.231*** (3.050)
Fetus Presentation Transverse	−317.582*** (6.015)	−338.577*** (12.743)
Fetus Presentation Ignored	−84.498*** (3.726)	−53.431*** (7.378)
Assisted Birth Nurse	39.578*** (1.276)	15.133** (7.404)
Assisted Birth Midwive	−0.475 (5.501)	8.017 (60.130)
Assisted Birth Others	−10.640** (5.286)	−52.142 (41.324)
Assisted Birth Ignored	−28.277* (14.839)	11.304 (35.357)
Birth Place Non Hospital		−14.708* (8.093)
−23.375* (14.174)		
Observations	4,185,558	812,305
R^2	0.019	0.013
Adjusted R^2	0.018	0.012
F Statistic	138,241.700***	22,761.500***

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the results for the 2SLS models using prenatal care delay as endogenous variable. Column 1 refers to the sample of births in SUS health units and column 2 to strictly private ones. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 32 – Heterogenous effects: Ethnicity - Month of Start

	Dependent Variable: Birth Weight	
	White	Black and Brown
	(1)	(2)
Month of Start	-97.581^{***} (4.682)	-66.188^{***} (3.042)
Previous Gestations	33.915^{***} (0.601)	24.805^{***} (0.427)
C-Section	80.772^{***} (1.165)	99.329^{***} (0.846)
Mother's Age	-1.537^{***} (0.128)	1.551^{***} (0.101)
Sex Ignored	-720.071^{***} (36.592)	-601.287^{***} (24.107)
Sex Male	111.388^{***} (0.824)	105.161^{***} (0.628)
Mother's education College	32.750^{***} (4.938)	32.209^{***} (2.594)
Mother's Education Primary	20.364^{***} (4.746)	29.768^{***} (2.131)
Mother's education High School	42.728^{***} (4.735)	47.147^{***} (2.237)
Mother's education Ignored	35.703^{***} (8.055)	29.609^{***} (3.915)
Mother's education None	-35.759^{**} (13.894)	-39.023^{***} (5.269)
Induced	84.412^{***} (1.244)	76.854^{***} (0.935)
Fetus Presentation Pelvic	-307.411^{***} (2.111)	-305.466^{***} (1.778)
Fetus Presentation Transverse	-384.593^{***} (8.805)	-285.611^{***} (7.198)
Fetus Presentation Ignored	-79.713^{***} (6.016)	-73.847^{***} (4.194)
Assisted Birth Nurse	40.332^{***} (2.585)	39.756^{***} (1.474)
Assisted Birth Midwive	-0.705 (17.523)	2.006 (6.007)
Assisted Birth Others	0.080 (15.010)	-11.378^{**} (5.788)
Assisted Birth Ignored	-39.672 (28.152)	-21.700 (16.676)
Birth Place Non Hospital		-26.834^{**} (13.229)
-12.980 (8.693)		
Observations	1,745,317	3,051,909
R^2	0.015	0.020
Adjusted R^2	0.013	0.018
F Statistic	$56,330.660^{***}$	$97,850.960^{***}$

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the results for the 2SLS models using prenatal care delay as endogenous variables. Column 1 refers to the sample of white mothers and column 2 to black and brown ones. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level.
$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

TABLE 33 – Heterogenous effects: Mother's Education - Month of Start

	Dependent Variable: Birth Weight			
	Primary	Secondary	High School	College
	(1)	(2)	(3)	(4)
Month of Start	−129.573*** (19.365)	−74.496*** (6.919)	−75.000*** (3.035)	−59.732*** (6.426)
Previous Gestations	21.108*** (1.946)	23.108*** (0.929)	28.159*** (0.453)	33.354*** (0.654)
C-Section	63.456*** (5.193)	85.423*** (1.776)	97.917*** (0.836)	86.015*** (1.566)
Mother's Age	−0.229 (0.473)	2.249*** (0.227)	0.999*** (0.100)	−3.190*** (0.147)
Sex Ignored	−816.684*** (122.274)	−511.334*** (44.583)	−632.782*** (24.057)	−626.296*** (52.338)
Sex Male	102.597*** (3.578)	105.082*** (1.239)	106.885*** (0.629)	111.736*** (1.065)
Race White	−14.631 (35.354)	7.871 (12.253)	−9.886* (5.386)	25.829*** (6.434)
Race Ignored	−3.310 (36.712)	12.831 (12.867)	−1.242 (5.863)	32.355*** (7.671)
Race Indigenous	−25.662 (37.431)	7.791 (14.193)	24.805*** (7.725)	81.461*** (17.065)
Race Brown	−13.661 (34.968)	11.204 (12.152)	−4.461 (5.357)	38.578*** (6.492)
Race Black	−30.868 (35.582)	−3.310 (12.375)	−15.103*** (5.494)	28.266*** (7.028)
Induced	91.201*** (5.847)	83.605*** (1.862)	80.616*** (0.912)	68.857*** (1.821)
Fetus Presentation Pelvic	−294.136*** (9.865)	−307.232*** (3.567)	−317.674*** (1.768)	−291.404*** (2.593)
Fetus Presentation Transverse	−150.551*** (33.247)	−280.668*** (13.506)	−317.952*** (7.316)	−394.012*** (11.191)
Fetus Presentation Ignored	−24.178 (24.694)	−87.010*** (8.448)	−79.572*** (4.479)	−80.618*** (6.790)
Assisted Birth Nurse	29.419*** (7.873)	37.843*** (2.832)	42.270*** (1.525)	30.057*** (4.297)
Assisted Birth Midwive	−4.484 (21.077)	−1.666 (10.173)	−0.318 (7.405)	−10.128 (25.680)
Assisted Birth Others	−44.949** (22.206)	−9.500 (9.971)	−9.949 (6.943)	−25.722 (20.786)
Assisted Birth Ignored	112.448 (78.254)	−53.716* (30.677)	−19.808 (17.922)	−8.696 (35.381)
Birth Place Non Hospital	−26.280* (15.969)	−8.950 (46.863)	−34.138** (17.004)	−9.943 (9.265)
Observations	114,341	804,839	3,016,577	1,012,588
R^2	0.008	0.019	0.018	0.021
Adjusted R^2	−0.020	0.014	0.017	0.018
F Statistic	2,671.617***	26,469.220***	98,827.070***	31,973.600***

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the results for the 2SLS models using prenatal care delay as endogenous variable. Column 1 refers to the sample of women with less than three years of education. In column 2, four to seven years of schooling. In column 3, eight to eleven years of schooling. In column 4, more than twelve years of schooling. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 34 – Heterogenous effects: Region of Birth - Month of Start

	Dependent Variable: Birth Weight				
	South	Southeast	MidWest	North	Northeast
	(1)	(2)	(3)	(4)	(5)
Month of Start	−116.374*** (7.489)	−122.695*** (3.910)	−85.950*** (7.691)	−50.187*** (8.122)	29.511*** (4.334)
Previous Gestations	40.577*** (1.312)	32.900*** (0.577)	19.994*** (0.809)	32.076*** (1.307)	14.658*** (0.552)
C-Section	75.080*** (1.916)	78.044*** (1.073)	76.246*** (2.043)	88.390*** (2.319)	135.311*** (1.257)
Mother's Age	−2.794*** (0.252)	−1.300*** (0.119)	1.052*** (0.213)	2.194*** (0.342)	4.407*** (0.139)
Sex Ignored	−691.567*** (72.593)	−638.502*** (31.078)	−529.952*** (60.175)	−537.023*** (77.985)	−580.794*** (31.185)
Sex Male	111.799*** (1.362)	108.887*** (0.792)	108.433*** (1.497)	102.922*** (1.528)	103.738*** (0.947)
Race White	16.995 (14.216)	7.181 (5.596)	−3.520 (9.922)	−30.120* (15.787)	−11.512 (8.542)
Race Ignored	16.878 (16.353)	9.394 (6.668)	9.668 (10.449)	−15.312 (16.838)	−1.969 (8.907)
Race Indigenous	70.083*** (21.589)	191.931*** (12.574)	44.867*** (13.739)	−47.216*** (16.647)	−27.923** (11.923)
Race Brown	22.702 (14.281)	13.397** (5.598)	13.394 (9.857)	−16.613 (15.609)	−4.102 (8.442)
Race Black	19.791 (14.587)	−3.738 (5.744)	15.271 (10.606)	−14.451 (16.464)	−4.821 (8.684)
Mother's education College		34.440*** (4.223)	−4.337 (7.604)	35.392*** (6.522)	54.484*** (3.531)
Mother's Education Primary		25.442*** (4.006)	6.283 (6.811)	32.809*** (4.664)	37.405*** (2.794)
Mother's education High School		43.971*** (4.017)	14.676** (6.971)	50.331*** (5.173)	68.091*** (2.953)
Mother's education Ignored		26.842*** (6.719)	5.794 (11.282)	42.527*** (8.577)	39.888*** (4.863)
Mother's education None		−24.427** (12.023)	3.639 (16.999)	−47.418*** (9.847)	−64.613*** (6.612)
Induced	106.799*** (1.917)	72.616*** (1.114)	69.410*** (2.367)	61.207*** (2.766)	74.738*** (1.542)
Fetus Presentation Pelvic	−308.437*** (3.401)	−331.513*** (2.151)	−240.957*** (4.173)	−220.558*** (4.491)	−317.071*** (2.634)
Fetus Presentation Transverse	−371.509*** (12.763)	−388.098*** (9.241)	−317.916*** (19.706)	−169.590*** (16.241)	−255.597*** (10.530)
Fetus Presentation Ignored	−71.452*** (10.337)	−96.138*** (6.228)	−98.708*** (7.095)	−20.618*** (7.910)	−80.619*** (7.752)
Assisted Birth Nurse	28.032*** (5.717)	60.371*** (1.915)	50.296*** (6.901)	8.208*** (3.150)	21.888*** (2.313)
Assisted Birth Midwive	−50.389 (38.934)	4.240 (28.206)	6.111 (51.781)	−0.136 (8.803)	−7.466 (7.325)
Assisted Birth Others	−84.316** (33.647)	−29.659* (16.675)	−29.721 (30.801)	−12.804 (8.170)	−12.583 (7.777)
Assisted Birth Ignored	79.820 (55.605)	−35.210 (28.250)	−22.776 (27.232)	−47.649 (35.495)	−25.350 (24.478)
Birth Place Non Hospital		−11.874 (22.892)	−29.384** (13.717)	12.069 (25.852)	−7.234 (19.944)
−17.552 (10.697)					
Observations	663,736	2,014,107	501,721	485,551	1,332,748
R^2	0.014	0.012	0.010	0.021	0.030
Adjusted R^2	0.013	0.011	0.009	0.020	0.029
F Statistic	22,557.860***	61,511.850***	13,159.960***	15,485.270***	50,003.990***

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the results for the 2SLS models using prenatal care delay as endogenous variable. Estimate 1 represent the results for the sample born in the southern region. Estimate 2 for the southeast. Estimate 3 for the midwest. Estimate 4 for the north. Estimate 5 for the northeast. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

TABLE 35 – Heterogenous effects: Genetic Anomaly - Month of Start

	Dependent Variable: Birth Weight	
	Normal	Genetic Anomaly
	(1)	(2)
Month of Start	−73.939*** (2.560)	−103.556*** (31.968)
Previous Gestations	28.958*** (0.359)	21.287*** (3.606)
C-Section	93.179*** (0.679)	126.507*** (9.461)
Mother's Age	0.365*** (0.081)	−2.052** (0.818)
Sex Ignored		−487.688*** (31.473)
Sex Male	107.810*** (0.493)	132.218*** (7.132)
Race White	4.837 (3.958)	−3.214 (45.086)
Race Ignored	11.547*** (4.362)	39.232 (52.777)
Race Indigenous	27.530*** (5.556)	119.405 (75.979)
Race Brown	11.198*** (3.948)	28.942 (45.164)
Race Black	−0.745 (4.071)	35.119 (47.157)
Mother's education College	34.107*** (2.173)	−41.033 (31.936)
Mother's Education Primary	27.894*** (1.895)	−34.952 (27.343)
Mother's education High School	47.781*** (1.961)	−16.590 (28.927)
Mother's education Ignored	31.498*** (3.391)	−74.530 (56.072)
Mother's education None	−35.759*** (4.671)	−121.895* (64.333)
Induced	79.164*** (0.739)	134.563*** (10.559)
Fetus Presentation Pelvic	−303.484*** (1.352)	−370.916*** (12.677)
Fetus Presentation Transverse	−315.189*** (5.503)	−577.356*** (59.418)
Fetus Presentation Ignored	−81.908*** (3.625)	−220.763*** (44.904)
Assisted Birth Nurse	38.207*** (1.258)	82.571*** (18.484)
Assisted Birth Midwive	−3.293 (5.471)	49.274 (116.733)
Assisted Birth Others	−11.884** (5.245)	69.243 (91.912)
Assisted Birth Ignored	−16.098 (14.599)	138.191 (216.824)
Birth Place Non Hospital		−9.437 (7.194)
−95.011 (58.773)		
Observations	4,859,663	44,726
R^2	0.018	0.026
Adjusted R^2	0.017	−0.025
F Statistic	154,447.300***	2,093.651***

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the results for the 2SLS models using prenatal care delay as endogenous variables. Column 1 refers to the sample of normal births and column 2 births with genetic anomaly. All estimates are being controlled for hospital and municipality fixed effects and maternal and infant covariates. The list of covariates include: schooling, age, marital status, race, number of dead children, number of prenatal visits, induced labor indicator, assisted birth status, fetus presentation and place of birth . The sample is restricted to births between 2015 and 2017. Robust standard errors (in parentheses) are clustered at the fixed effects level. *p<0.1; **p<0.05; ***p<0.01

3 THE IMPACT OF C-SECTIONS ON BREECH PREGNANCIES IN BRAZIL: AN ASSESSMENT USING PROPENSITY SCORES

3.1 ABSTRACT

This study examined the impact of having a cesarean section (C-Section) on newborns born from breech pregnancies using a sample of 28 thousands births from the Brazillian Unique Health System (SUS). An inverse probability of treatment weighting method was used to measure the c-section impact on the infant´s APGAR scores and mortality in the first year of life, addressing the self selection bias inherent in this setting. Our findings are that, for breech babies, having a C-Section decreases the probability of having low APGAR scores [odds ratio (OR): 0.577 ; IR: 0.521-0.640] and death in the first year of life [odds ratio (OR): 0.617 ; IR: 0.503-0.766]. There is no evidence of impact in the probability of having low weight birth (<2500g). Given these findings, we argue that health policymakers should take a contextual approach when reducing global cesarean rates in order to mitigate potential risks for newborns of high risk pregnancies.

Keywords: Cesarean Section; Newborn health; Breech Pregnancies; Propensity Scores; Brazil

3.2 INTRODUCTION

Health policymakers are increasingly adopting evidence based frameworks to support policy decisions (COOKSON, 2005). A particular decision that is illustrative of the challenges and tradeoffs that health policymakers and practitioners face is related to the mode of delivery in breech pregnancies - when the baby is born in an inverted position, because it is sitting inside the belly during birth. A breech birth is more risky than a normal birth because there is a possibility that the baby could become stuck in the birth canal or the umbilical cord could become twisted or compressed during delivery, which could lead to decreased oxygen supply, increasing the risk of damage to the baby's body and brain (FISCHER-RASMUSSEN; TROLLE, 1967).

The most important research on C-Sections and breech pregnancies was the 'Term Breech Trial' (TBT) (HANNAH et al., 2000). This randomized trial study influenced policy making and health practices worldwide by showing a positive health impact of C-Sections in breech pregnancies. After the TBT study, the rate of C-Sections in breech presentations began to increase across different countries (RIETBERG; ELFERINK-STINKENS; VISSER, 2005). However, the results from the 'Term Breech Trial' study have been under intense scrutiny and criticism (GLEZERMAN, 2006; LAWSON, 2012). Nonetheless, recent observational studies have showed positive health results for C-sections in breech births (JENSEN; WÜST, 2015;

MÜHLRAD, 2018). In sum, although there are studies favoring delivery trough C-Sections, there is no established consensus on the best mode of delivery in breech pregnancies and this poses a challenge to health policymakers.

This paper´s contribution is to assess, for the first time, the impact of C-Sections on the health of breech babies in Brazil - a highly heterogeneous middle income country. We use nationally representative birth and mortality microdata of 28 thousand breech births from the Brazillian Unified Health System (SUS). The empirical strategy is to use a inverse probability of treatment weighting (IPTW) framework to deal with treatment self selection and regional heterogeneity in the relationship between birth by C-section and infant´s health.

The findings from the weighted propensity score models are that C-Sections have a significant negative effect on the probablility of breech babies having low APGAR scores and dying in their first year of life. There is no significant effect effect on the the probability of low weight birth (<2500g). Our findings can be interpreted as a cautionary tale. Even though there is a public policy rationale to reduce the widespread adoption of C-Sections in Brazil, the reduction should be made on a contextual basis, accounting for the specific necessities of mothers and infants.

3.3 LITERATURE REVIEW

3.3.1 Demand for C-Section in Brazil

The World Health Organization (WHO) recommendation for the total proportion of cesarean sections (C-Sections) is 15 percent of all births (BETRAN et al., 2016). There are significant discrepancies in a woman's access to cesarean sections depending on where she lives in the world. About 8% of women gave birth by C-Sections in the least developed countries, with just 5% in sub-Saharan Africa. On the other hand, in Latin America and the Caribbean, the rates reach four in 10 (43%) births. In five countries (Dominican Republic, Brazil, Cyprus, Egypt, and Turkey), C-Sections now outnumber normal deliveries (BETRAN et al., 2021).

Several studies in the public health and epidemiology literature help understand the context and the reasons behind the high rate of c-sections in Brazil. A study accompanying four birth cohorts in Pelotas - a city in the south of Brazil - has seen the proportion of c-sections increase from 27.6% in 1982 to 65.1% in 2015, with the rate in the private sector reaching 93.9% (BARROS et al., 2019). A nationwide observational study using Unified Health System (SUS) data, documents a increase in c-sections from 37.9% in 2000 to 53.9% in 2011 (BARROS et al., 2015). The most important socioeconomic characteristics associated with c-sections were being white, older, having more education, and having their first pregnancy [1].

[1] However, the relative composition of the socioeconomic groups has changed during the period. Between 1991 and 2006, the odds of doing a c-section diminished for wealthy families while it increased for families with low socioeconomic status (RAIFMAN; CUNHA; CASTRO, 2014). One possible reason is that low-income women tend to seek birth by c-section for fear of low-quality care due to social status (BÉHAGUE; VICTORA;

For a sample of women in two private sector hospitals in Rio de Janeiro, (DIAS et al., 2008) finds that, although 70% of them did not report an initial preference for cesarean section, 90% had one. Irrespective of the initial desire of the mother, the health services pregnancy process resulted in a c-section as the final route of delivery. A more recent study with a nationwide hospital cohort has found similar proportions for mothers' birth preferences, 27% for c-sections and 73% for vaginal births (DOMINGUES et al., 2014). In both studies, mothers state that fear of pain is the major motive for preferring a c-section.In a particular sample of teenage mothers, the most important factor associated with c-sections was the perception that it is a safer mode of delivery (GAMA et al., 2014). We can argue that the high proportion of c- sections cannot be attributed to a single factor such as doctor's incentives or mother's preferences (FAÚNDES; CECATTI, 1993). Instead, it should be understood as a complex, multifaceted situation where the interaction of social groups realities, institutional factors, cultural preferences, and economic incentives are driving the high prevalence of c-sections in Brazil (MCCALLUM, 2005).

3.3.2 Impact of C-Sections: Empirical Strategies

Identifying the causal effect of a C-Section on a baby's health is challenging. There might be selection effects that influence both variables: such as richer mothers, who have higher health endowments, having more C-Sections than poorer ones. On the other hand, the group of mothers with higher risk pregnancies also might have more C-sections than lower risk ones. Overall, there is an understanding that the relationship between the infant's health and the choice of delivery is endogenous. Therefore, to understand this relationship without bias, there is a need to establish an identification strategy to deal with the endogeneity between these two variables. Recently, many studies have tried to exploit natural experiments, instrumental variable strategies or matching methods to address this research problem.

In Spain, (COSTA-RAMÓN et al., 2018) uses differences in the time of birth as an instrumental variable. Finds that c-sections without a medical recommendation have a negative impact on the newborn APGAR scores but do not have an effect on mortality. (BORRA; GONZÁLEZ; SEVILLA, 2016) explores a quasi-natural experiment when the Spanish government stopped paying a monetary bonus for babies born. They find no increase in health problems at birth but an increase in the probability of the newborn being hospitalized for respiratory problems in the first months of life. In the United States, (CARD; FENIZIA; SILVER, 2019) explores the distance between hospitals to identify the impact of c-sections on an infant's health. In particular, they use the mother's proximity to hospitals with high or low cesarean section rates as an instrumental variable. In the short term, the impact of c-sections seems to be positive on the Apgar score and neonatal mortality. However, they find evidence that babies born from c-section are more likely to develop respiratory problems in the long run.

BARROS, 2002)

(SCHULKIND; SHAPIRO, 2014) uses a natural experiment of tax rewards to births that occur slightly before new year's eve to assess the impact of elective c-sections on a newborn's health. They find that even small anticipations in birth time can cause lower APGAR scores and birth weight for babies.

Few studies aim to understand the impact of C-Sections in the infant´s health in Brazil. Using a logistic regression model for different cohorts of mothers, a study do not find evidence for an increased probability of long-term respiratory conditions such as asthma in c-section babies (MENEZES et al., 2011). A research using survival analysis weighted by propensity scores to assess the impact of c-section in the infant's health of different Robson groups [2], finds an increase in infant mortality for groups where c-sections are typically not recommended but a decrease for groups where they are - breech pregnancies included (PAIXAO et al., 2021b).

3.3.3 C-Sections and Breech Pregnancies

The 'Term Breech Trial' (TBT) is the most cited and influential paper on breech pregnancies (HANNAH et al., 2000). Spanning 26 countries, the study consisted in randomly allocating mothers with breech babies to planned vaginal or planned C-Section. They found that C-Sections were better than vaginal births for breech babies in terms of infant mortality and morbidity. After its publication, the TBT influenced the choice of delivery in breech pregnancies worldwide, with a substantial increase in planned C-sections (RIETBERG; ELFERINK-STINKENS; VISSER, 2005). However, there are concerns regarding the proper interpretation of the TBT results - such as practice differences between countries or the lack of experienced obstetricians in many observations (TURNER, 2006; GLEZERMAN, 2006).

Later studies have tried to assess the impact of C-Sections on breech babies, taking into account the impact of the TBT study. Jensen and Wust (2015) explore the impact that the 'Term Breech Trial' (TBT) in cesarean rates in Denmark. They use a regression discontinuity framework that explores the 'information shock' in obstetricians and patients incentives caused by the TBT study. They find that having a c-section diminishes the probability of breech babies having low APGAR scores and hospitalizations in the first year and do not find significant mortality results. Muhlrad (2018) explores the same 'information shock' to use a regression discontinuity in Sweden. The author also finds that C-Sections improve newborn´s health in the short therm and also during childhood.

On the other hand, a Finnish national cohort study found that C-Sections in breech pregnancies increase the probability of worse outcomes (such as low APGAR scores and intensive unit care admission) in subsequent deliveries (MACHAREY et al., 2020). Also in Finland, a retrospective register-based study found that vaginal breech births at 32 to 36 weeks of gestation

[2] The physician Michael Robson created the Robson classification in 2001 with the aim of prospectively identify groups of women clinically relevant, where there are differences in cesarean section rates and thus allowing comparisons in the same institution over time or between different institutions (BETRAN et al., 2014)

do not increase morbidity or infant mortality relative to C-Sections (TOIJONEN et al., 2022). A case control study in Portugal, using a small sample of 26 treated observations finds that vaginal deliveries are as safe as C-Sections for breech pregnancies when there is the presence of skilled obstetricians (VALENTE; AFONSO; CLODE, 2020). That is, they do not find any increase in mortality for vaginal deliveries. A Brazilian study reviews the literature on breech pregnancies and indicates that ,while there is no consensus on the best mode of delivery in these cases, planned vaginal births can be a safe option in specific contexts - such as when they are performed by experienced obstetricians (SIMÕES et al., 2015).

3.4 DATA

Our empirical strategy requires birth characteristics and infant mortality data. To this end, we use microdata from different Ministry of Health (MS) sets: i) birth characteristics from the vital statistical records of the Live Birth Information System (SINASC); ii) mortality from the Mortality Information System (SIM); and iii) to support our analysis, hospital characteristics from the National Register of Health Establishments (CNES).

The Live Birth Information System (SINASC) was implemented by the Brazilian Ministry of Health to systematically record information on live births at the national level. The SINASC is based on the Declaration of Live Birth (DN) unique identifier, including data on mother and newborn characteristics. Health professionals or traditional midwives fill in the DNs. Every newborn has an identification number in the live birth information system - the DN number - as well as the health unit where the birth procedure took place - the CNES number.

The National Registry of Health Establishments (CNES) is the official information system for registering information from all health establishments, regardless of their legal nature or whether they are part of the Unified Health System (SUS). This is the official registry of the Ministry of Health (MS) regarding the reality of installed capacity and health care workforce in Brazil in public or private health establishments, with or without SUS agreement.

In Brazil, it is required that all municipalities pass death records information to the state and federal governments. Then, the process of collecting, storing, and managing death records is consolidated in the Mortality Information System (SIM). These records have information on causes of death, date of death, and socioeconomic characteristics of virtually all registered deaths in Brazil. The information on infant mortality (below one year of age) can be identified using the DN number.

To build our dataset, we begin using the 2017 Live Birth System (SINASC) with a total of 2.97 million observations uniquely identified by the DN number. Each birth in the SINASC also has a unique identifier for the associated health unit where it occurred, the CNES number. We then match the information from the SINASC with the National Register of Health Establishments using the unique CNES number identifier. The CNES data allow us to identify

TABLE 36 – Variables Description

Variable Name	Definition	Type	Source
DN	Infant ID	Numerical	Livebirth Dataset (SINASC)
CNES	Hospital ID	Numerical	Livebirth Dataset (SINASC)
Birth Location	Place of Birth	Categorical	Livebirth Dataset (SINASC)
Mother Age	Mother´s age in years	Numerical	Livebirth Dataset (SINASC)
Marital Status	Marital Status	Categorical	Livebirth Dataset (SINASC)
Schooling	Mother´s Education in Years of Education	Numerical	Livebirth Dataset (SINASC)
Live Children	Number of living children	Numerical	Livebirth Dataset (SINASC)
Dead Children	Number of dead children	Numerical	Livebirth Dataset (SINASC)
Gestational Weeks	Gestational Weeks	Numerical	Livebirth Dataset (SINASC)
Parity	Type of Pregnancy (Unique; Double; Triple)	Categorical	Livebirth Dataset (SINASC)
C Section	Type of delivery: Vaginal or C-Section	Categorical	Livebirth Dataset (SINASC)
Prenatal Visits	Number of pre-natal care visits	Numerical	Livebirth Dataset (SINASC)
Sex	Infant Sex	Categorical	Livebirth Dataset (SINASC)
APGAR1	1st minute APGAR	Numerical	Livebirth Dataset (SINASC)
APGAR5	Fith minute APGAR	Numerical	Livebirth Dataset (SINASC)
Race	Race/Ethnicity	Categorical	Livebirth Dataset (SINASC)
Weight	birth weight in grams	Numerical	Livebirth Dataset (SINASC)
Anomaly	Genetic Anomaly	Categorical	Livebirth Dataset (SINASC)
Mother Ethnicity	Mother´s Race/Ethnicity	Categorical	Livebirth Dataset (SINASC)
Previous Gestations	Number of Previous Gestations	Numerical	Livebirth Dataset (SINASC)
Vaginal Births	Number of Vaginal Births	Numerical	Livebirth Dataset (SINASC)
Cesarean Births	Number of c-sections	Numerical	Livebirth Dataset (SINASC)
Death	If infant died in the first year of life	Binary	Mortality Information System (SIM)

Source: Prepared by the authors using Unified Health System (SUS) data.
Note: The table describes the variable name that we adopted in our estimations, the definition of each variable based on SUS data, the variable type and the corresponding data source from the SUS.

which health unit is associated with the Unique Health System (SUS) and which are not.

The third step is to match these two datasets with the Mortality Information System (SIM) using the DN number. The SIM data allows us to identify infants born in 2017 and who have died in the same year. The matched dataset has information on hospital, mortality, and live birth characteristics. We then remove missing data and incomplete information in the matched dataset. The result is a sample with around one million births from 2017. Table 36 defines the most relevant variables in our sample, as well as their type and data source.

Table 37 describes the summary statistics for breech pregnancies (first column) and for all other pregnancies (second column). The summary statistics table indicate that there is important mean differences between breech and all other pregnancies. Indeed, mean APGAR scores are lower and mean birthweight is more than 150 grams lower for breech babies. Our focus of analysis in this work will be the breech births sample, with around 28 thousand observations.

TABLE 37 – Summary Statistics

Variable Name	Breech Pregnancies	All Other Pregnancies
Mother Age	27.797	26.632
	(6.890)	(6.705)
Live Children	0.924	1.022
	(1.326)	(1.436)
Number of dead children	0.283	0.242
	(0.710)	(0.859)
Gestational Weeks	37.663	38.535
	(3.168)	(2.185)
Prenatal Visits	8.462	8.363
	(7.887)	(6.853)
APGAR1	7.927	8.441
	(2.454)	(2.722)
APGAR5	9.123	9.411
	(2.177)	(2.331)
Birth Weight	3002.825	3196.054
	(736.021)	(553.927)
Previous Gestations	1.156	1.232
	(1.496)	(1.579)
Vaginal Births	0.565	0.705
	(1.284)	(1.464)
Cesarean Births	0.418	0.369
	(0.905)	(1.015)
Observations	28693	1171307

Source: Prepared by the authors using Unified Health System (SUS) data.
Note: The table show the average and the standard deviation (in parentheses) for the numeric variables of our sample.

3.5 EMPIRICAL STRATEGY

This study objective is to estimate the impact of cesarean sections in the health of breech babies. Since it is a observational setting, a key challenge in estimating the impact of C-Sections on the infant's health is selection bias. That is, mother´s self-selecting to have cesarean births in such a way as to change the observable characteristics between 'treated' and 'control' groups: one group is oversampled relative to the hypothetical sample from a randomized experiment [3]. Therefore, women self-selecting into giving birth by C-sections violate the random attribution mechanism, which is a key condition to infer causality.

A standard modelling strategy to deal with selection bias is the inverse probability of treatment weighting (IPTW)(AUSTIN; STUART, 2015). IPTW creates a weighted population where treatment assignment no longer depends on the covariates. In the original population, some mothers are more likely to get treated than others, based on their observable characteristics.

[3] In causal analysis, the ideal setting to infer the impact of a variable on another is to do a experiment with a random assignment mechanism. The thinking behind the random assignment is to randomize treatment to groups with essentially equivalent characteristics. Thus, any effect observed between treated groups may be linked to the treatment effect and not to the different characteristics of the individuals in the group (KRAUSE; HOWARD, 2003).

In the weighted population, every mother is equally likely to be treated despite the differences in their covariates. That way, there is is no confounding in the weighted population.

Mother's with breech pregancies (Robson groups 6 and 7 [4]) are divided between treatment and control group based on their type of birth. Mothers who did vaginal births are in the control group ($T_i = 0$, for woman i) and mothers who did C-Sections are in the treatment group ($T_i = 1$). First, the IPTW approach consists of estimating the probability of having a C-Section (T) conditional on the observed covariates $(x_i) : Pr(x_i) = Pr(T = 1|x_i)$. And second, to divide each observation in the sample by their propensity score $\frac{1}{n}\sum_i \frac{Ty}{Pr}$. The result is a sample where mothers are weighted by their probabilities of having a c-section(their propensities scores).

A key assumption in weighting strategies is that a common support exists between treated and control groups. That way, to test the quality of the IPTW procedure, it is typical to compare the covariates' standard mean differences between treated and control groups. Finally, a set of logit model will be estimated to measure how C-Section's impact the probability of infant's death, low birth weight (<2500g) and low APGAR scores (below 7).

3.6 RESULTS

Figure 18 shows the characteristics of women in the treatment and control group. In summary, essentially all variables have standard mean differences lesser than 0.1. The difference being bounded below 0.1 indicates that the observable characteristics of the individuals in these two different samples are similar (STUART, 2010). There is a "common support" between 'treated' and 'control' groups—a key element for the identification of effects in our empirical strategy.

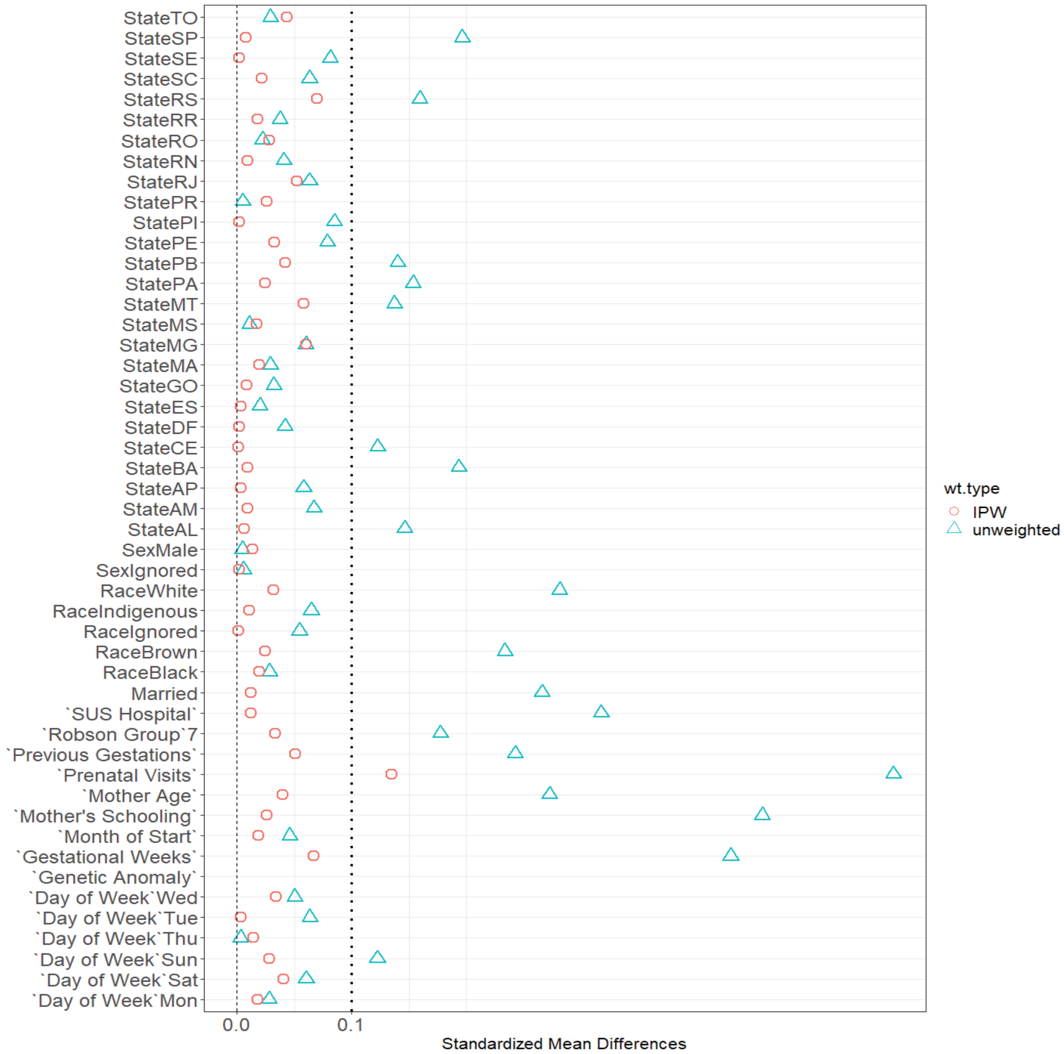
In more detail, table 38 show that after weighting, women are similar in terms of race, marital status, and schooling. The number of prenatal care visits is close to 6.5 for both samples. The mother's age is close to 26. Gestational weeks is similar for both groups. The balance between samples for gestational weeks is especially important because they can have a direct influence on the baby's health [5]. The balance in days of the week is also an important result since there is well documented evidence of the existence of selection effects in choosing specific birth days [6].

[4] Group 6 contains all breech births from first pregnancies and group 7 all breech births from mothers who already had previous births - including mothers who had previous C-Sections (NAKAMURA-PEREIRA et al., 2016)

[5] In São Paulo, Brazil most populated state, there is evidence of a one-week reduction in the gestational age for those born by cesarean section in the private sector (DINIZ et al., 2016). The covariate balance between both samples is indicative that this effect is being accounted for and is not cofounding our results.

[6] The influence of medical convenience, idiosyncratic choices, or financial incentives in changing birth dates has been well discussed in the literature. Using data from military hospitals in the USA (III, 1996) has shown the lower probability of cesareans taking place on weekends and the higher probability that they occur between 6 pm and 12 am. Or (LO, 2003) that indicates a reduction in births on specific dates considered inauspicious in popular beliefs. (SPINOLA, 2016) based on the analysis of working days between holidays,

FIGURE 17 – Covariate Balance - IPTW



Source: Prepared by the authors using Unified Health System (SUS) data.

Notes: The figure presents the covariate balance between treatment and control groups after weighting by propensity scores.

Table 39 shows the impact of C-Sections on the infant´s health measures for breech babies. The first column presents the impact of C-Sections in the probability of death in the first year of life. It is a strongly significant negative result [odds ratio (OR): 0.617 ; IR: 0.503-0.766]. The second column, indicate a reduction in the probability of having low APGAR1 scores for breech babies born trough C-Sections [odds ratio (OR): 0.577 ; IR: 0.521-0.640]. Results are more substantial for the probability of low APGAR5 [odds ratio (OR): 0.339 ; IR: 0.285-0.405]. Finally, column (4) shows the impact in the probability of the infant having low weight (<2500g) [odds ratio (OR): 1.035 ; IR: 0.931-1.152] . There is no statistically significant result for this variable.

finds evidence of a convenience effect operating from physicians in the Brazilian public sector to shift these births to working days before or after the holiday. Finally, (GRUBER; KIM; MAYZLIN, 1999), and (GRANT, 2009) explore differences between the Medicaid and Medicare programs to infer how financial incentives influence the decision to perform a c-section.

TABLE 38 – Weighted Sample: Covariate Balance

Variable	Treated Group (Mean)	Control Group (Mean)	Standard Mean Difference
Distance	0.18	0.17	0.02
Race	0.32	0.29	0.06
Mother´s Schooling	5.04	5.02	0.01
Marital Status	0.28	0.27	0.01
Mother Age	26.60	26.59	0.01
Robson Group 6	0.34	0.32	0.05
Robson Group 7	0.65	0.67	0.01
Prenatal Care Visits	8.46	8.04	0.14
Prenatal Care Start	2.39	2.37	0.02
SUS Hospital	0.91	0.90	0.03
Male Infant	0.48	0.49	-0.02
Previous Gestations	1.42	1.461	-0.02
Gestational Weeks	36.05	36.17	-0.02
Race Black	0.05	0.06	-0.02
Race Brown	0.58	0.58	0.01
Race White	0.31	0.32	-0.07
Race Indigenous	0.008	0.009	-0.004
Day of Week Mon	0.155	0.149	0.001
Day of Week Tue	0.137	0.137	- 0.002
Day of Week Wed	0.147	0.146	0.001
Day of Week Thu	0.149	0.151	-0.006
Day of Week Sat	0.134	0.138	-0.011
Day of Week Sun	0.133	0.132	0.001
Number of Observations	26081	2612	28693

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows means for a set of maternal and pregnancy characteristics by treatment status and the standard mean difference for two groups: women who had C-Sections and women who had vaginal births. Two controls, state of birth and day of the week, were omitted but are shown in the appendix. The sample is restricted to single births, without genetic anomalies and with babies in breech position.

3.7 DISCUSSION

This paper is the first observational study to assess the impact of C-Sections in babies born from breech pregnancies using Brazillian microdata. There exists broader studies analyzing the impact of C-Sections on infant´s health in Brazil (PAIXAO et al., 2021b). However, we argue that there is value in focusing in breech pregnancies, because there is a not a established consensus in the literature regarding the best mode of delivery in these cases (SIMÕES et al., 2015)

Our main findings are that cesarean sections have a positive short-term result for breech babies health outcomes - decreasing the likelihood of death and low APGAR scores. We argue that these findings are robust to selection into treatment by a substantial degree since the inverse probability of treatment weighting (IPTW) framework creates a common support between treatment and control groups (AUSTIN; STUART, 2017; BISHOP; LEITE; SNYDER,

TABLE 39 – Treatment Effects - Weighted Data

	Dependent Variable			
	Infant Health			
	Death	Low APGAR1	Low APGAR5	Low Weight
	(1)	(2)	(3)	(4)
Treatment (C-Section)	-0.482***	-0.550***	-1.082***	0.034
	(0.107)	(0.052)	(0.089)	(0.054)
Observations	28693	28693	28693	28693
Log Likelihood	-3,135.63	-10,599.8	- 3,169.62	-12,607.8
Akaike Information Criteria	6,275.27	21,203.74	6,343.24	25,219.67

Source: Prepared by the authors using Unique Health System (SUS) data.
Notes: The table shows the estimates of the effect of having a C-Section on health measures of breech babies. The endogenous variable, an indicator of cesarean delivery, is instrumented with a dummy variable equal to one for deliveries between 8:00 am and 12:00 pm and 2:00 pm and 6:00 pm (Brazilian business hours). Outcome 1 is death in the first year of life. Outcome 2 and 3 are low APGAR1 and 5 scores (APGAR below 7) and outcome 4 is low weight (<2500g). All models are using IPTW weights controlled for birth differences in days of the week and states, maternal and infant characteristics. Maternal characteristics are: schooling, age, marital status, race, previous gestations, number of prenatal visits and month of prenatal start. Baby characteristics include: sex and gestational weeks. The sample is restricted to single births, without genetic anomalies and for babies in breech position. Standard errors (in parentheses).*p<0.1; **p<0.05; ***p<0.01

2018). These results are consistent with the findings of the Term Breech Trial (HANNAH et al., 2000) randomized experiment. As well as later observational studies (JENSEN; WÜST, 2015) (MÜHLRAD, 2018) that use regression discontinuity frameworks.

Our findings of cesarean sections reducing the risk of the infant´s having a low APGAR score in breech pregnancies are important because these scores are relevant as an indicator of an infant´s health. Indeed, low APGAR scores have been associated with an increased risk of infant mortality (ILIODROMITI et al., 2014), neonate mortality (LEE; SUBEH; GOULD, 2010), cerebral palsy (LIE; GRØHOLT; ESKILD, 2010), intellectual disabilities and autism (MODABBERNIA et al., 2019).

However, it is important to qualify that we do not have information on the long-run health of babies in our sample. Therefore we cannot assess if any future health complications are arising from the c-section such as those found in Card et al (2019). We also cannot evaluate the impact of C-Sections in subsequent deliveries (MACHAREY et al., 2020). Also an important caveat in our findings is that variations in the quality of medical care can impact the infant´s health results (COSTA-RAMÓN et al., 2018). We do not have a way to assess the quality of medical care. Not being able to control for the presence of experient obstetricians is a limitation, because they are an important factor to vaginal delivery safety in breech births (SIMÕES et al., 2015; VALENTE; AFONSO; CLODE, 2020)

Finally, our study restricted attention to the impact of c-sections on the infant´s outcomes and have not discussed the impact on the mother´s health. However, we point out that in Brazil, the high proportion of c-sections has been associated with maternal near-miss [7]

[7] Maternal near-miss happens when mothers almost die during the labor process. It is a metric that is used to measure the quality of obstetric care (LOTUFO et al., 2012)

(MENEZES et al., 2011). There is also more robust evidence of adverse health outcomes for mothers who had non-recommended c-sections (WANG et al., 2010) (SOUZA et al., 2010).

3.8 FINAL REMARKS AND PUBLIC POLICY IMPLICATIONS

There is an increasing public debate on the implications of the high proportions of cesarean sections in Brazil. However, the choice of delivery should be sensitive to the context in which the mother, the baby and the health providers are inserted. In particular, the best choice of delivery for breech pregnancies is a theme with mixed results in the health literature.

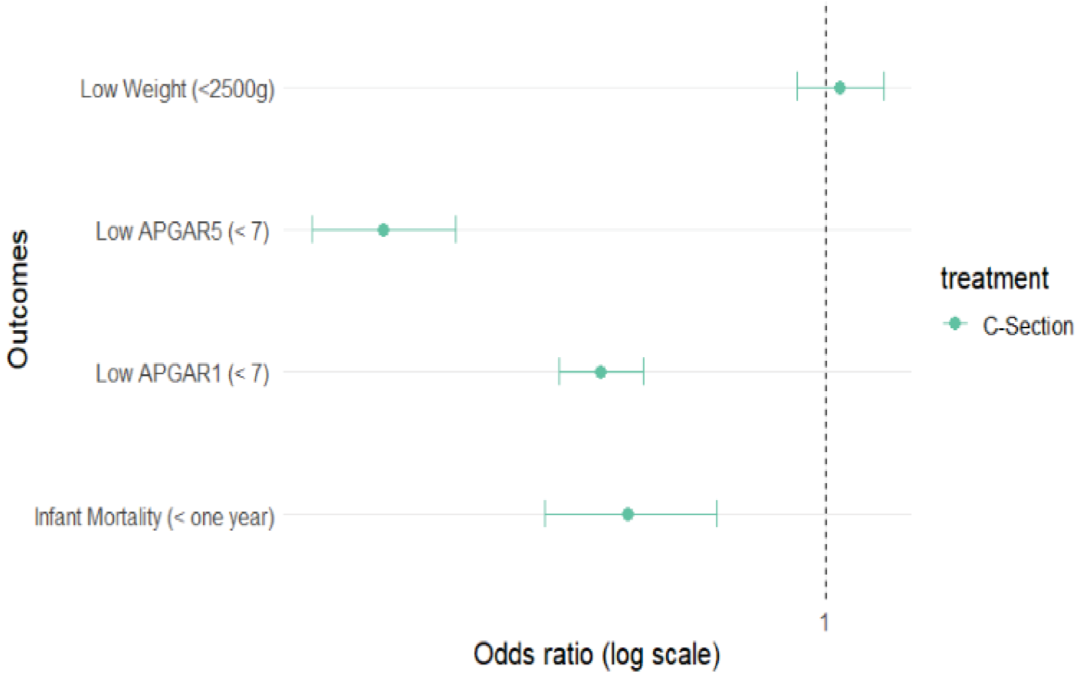
This research contributes to this discussion using a nationally representative microdata from the Brazillian Ministry of Health and a empirical strategy that deals with selection bias. That is, using an inverse probability of treatment weighting (IPTW) strategy, our main findings are that c-sections can decrease the probability of breech babies having low APGAR scores and death in the first year of life. We do not find evidence for decrease in the odds of low weight for breech babies born by c-section.

Health public sector officials should focus on developing a policy framework that can educate the population about the risks involved in different types of delivery to empower them to make informed decisions. Also, a well-crafted regulation effort to provide incentives to support evidence based policy making in the private sector and public sector can be an important measure. We argue that given our findings and the most recent literature, these measures can improve neonatal outcomes in Brazil.

Finally, we argue that there is a need for rigorous new studies that explore other associations between C-Section and birth outcomes. They may highlight relevant aspects and pathways through which the choice of delivery affects infants´ health in Brazil. In particular, studies that shed light on the long-term effects of C-Sections on babies from breech pregnancies are necessary to understand the public health implications of different modes of deliveries in high risk pregnancies.

3.9 APPENDIX

FIGURE 18 – Odds Ratio - Selected Outcomes



Source: Prepared by the authors using Unified Health System (SUS) data.

Notes: The figure presents the odds ratio (log scale) for the selected outcomes of newborn mortality, low weight (<2500g), low APGAR1 and 5. The treatment variable is having a C-Section instead of a vaginal delivery.

REFERENCES

ABATE, M. G.; ANGAW, D. A.; SHAWENO, T. Proximate determinants of infant mortality in ethiopia, 2016 ethiopian demographic and health surveys: results from a survival analysis. *Archives of Public Health*, BioMed Central, v. 78, n. 1, p. 1–10, 2020. Citado na página 22.

ALI, M.; SALEHNEJAD, R.; MANSUR, M. Hospital heterogeneity: what drives the quality of health care. *The European Journal of Health Economics*, Springer, v. 19, n. 3, p. 385–408, 2018. Citado na página 55.

ALIN, A. Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*, Wiley Online Library, v. 2, n. 3, p. 370–374, 2010. Citado na página 41.

ANGRIST, J. D.; KRUEGER, A. B. Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic perspectives*, v. 15, n. 4, p. 69–85, 2001. Citado na página 63.

ARAÚJO, T. V. B. de et al. Association between microcephaly, zika virus infection, and other risk factors in brazil: final report of a case-control study. *The Lancet Infectious Diseases*, Elsevier, v. 18, n. 3, p. 328–336, 2018. Citado na página 62.

ASSEMBLY, G. Sustainable development goals. *SDGs Transform Our World*, v. 2030, 2015. Citado na página 14.

ATHEY, S. Beyond prediction: Using big data for policy problems. *Science*, American Association for the Advancement of Science, v. 355, n. 6324, p. 483–485, 2017. Citado na página 46.

AUSTIN, P. C.; STUART, E. A. Moving towards best practice when using inverse probability of treatment weighting (iptw) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in medicine*, Wiley Online Library, v. 34, n. 28, p. 3661–3679, 2015. Citado na página 93.

AUSTIN, P. C.; STUART, E. A. The performance of inverse probability of treatment weighting and full matching on the propensity score in the presence of model misspecification when estimating the effect of treatment on survival outcomes. *Statistical methods in medical research*, SAGE Publications Sage UK: London, England, v. 26, n. 4, p. 1654–1670, 2017. Citado 2 vezes nas páginas 96 e 97.

AVDIC, D.; LUNDBORG, P.; VIKSTRÖM, J. Estimating returns to hospital volume: Evidence from advanced cancer surgery. *Journal of health economics*, Elsevier, v. 63, p. 81–99, 2019. Citado na página 55.

AWITI, J. O. A multilevel analysis of prenatal care and birth weight in kenya. *Health economics review*, BioMed Central, v. 4, n. 1, p. 1–16, 2014. Citado 3 vezes nas páginas 49, 56 e 57.

BANK, W. *World Development Report 2021: Data for Better Lives*. [S.l.]: The World Bank, 2021. Citado 2 vezes nas páginas 15 e 20.

BARNWAL, A.; CHO, H.; HOCKING, T. D. Survival regression with accelerated failure time model in xgboost. *arXiv preprint arXiv:2006.04920*, 2020. Citado na página 30.

BARROS, A. J. et al. Antenatal care and caesarean sections: trends and inequalities in four population-based birth cohorts in pelotas, brazil, 1982–2015. *International journal of epidemiology*, Oxford University Press, v. 48, n. Supplement_1, p. i37–i45, 2019. Citado na página 88.

BARROS, F. C. et al. Cesarean sections in brazil: will they ever stop increasing? *Revista Panamericana de Salud Pública*, SciELO Public Health, v. 38, p. 217–225, 2015. Citado na página 88.

BAȘ, E. K. et al. Maternal characteristics and obstetric and neonatal outcomes of singleton pregnancies among adolescents. *Medical science monitor: international medical journal of experimental and clinical research*, International Scientific Information, Inc., v. 26, p. e919922–1, 2020. Citado na página 43.

BATISTA, A. F. et al. Neonatal mortality prediction with routinely collected data: a machine learning approach. *BMC pediatrics*, BioMed Central, v. 21, n. 1, p. 1–6, 2021. Citado na página 16.

BAYINDIR, E. E. Hospital ownership type and treatment choices. *Journal of Health Economics*, Elsevier, v. 31, n. 2, p. 359–370, 2012. Citado na página 55.

BÉHAGUE, D. P.; VICTORA, C. G.; BARROS, F. C. Consumer demand for caesarean sections in brazil: informed decision making, patient choice, or social inequality? a population based birth cohort study linking ethnographic and epidemiological methods. *Bmj*, British Medical Journal Publishing Group, v. 324, n. 7343, p. 942, 2002. Citado na página 89.

BELUZO, C. E. et al. Machine learning to predict neonatal mortality using public health data from são paulo-brazil. *medRxiv*, Cold Spring Harbor Laboratory Press, 2020. Citado na página 16.

BETRAN, A. et al. Who statement on caesarean section rates. *BJOG: An International Journal of Obstetrics & Gynaecology*, Wiley Online Library, v. 123, n. 5, p. 667–670, 2016. Citado na página 88.

BETRAN, A. P. et al. A systematic review of the robson classification for caesarean section: what works, doesn't work and how to improve it. *PloS one*, Public Library of Science San Francisco, USA, v. 9, n. 6, p. e97769, 2014. Citado na página 90.

BETRAN, A. P. et al. Trends and projections of caesarean section rates: global and regional estimates. *BMJ Global Health*, BMJ Specialist Journals, v. 6, n. 6, p. e005671, 2021. Citado na página 88.

BISHOP, C. D.; LEITE, W. L.; SNYDER, P. A. Using propensity score weighting to reduce selection bias in large-scale data sets. *Journal of Early Intervention*, Sage Publications Sage CA: Los Angeles, CA, v. 40, n. 4, p. 347–362, 2018. Citado 2 vezes nas páginas 96 e 97.

BISHOP, C. M. Pattern recognition. *Machine learning*, v. 128, n. 9, 2006. Citado 2 vezes nas páginas 27 e 28.

BLAND, J. M.; ALTMAN, D. G. Survival probabilities (the kaplan-meier method). *Bmj*, British Medical Journal Publishing Group, v. 317, n. 7172, p. 1572–1580, 1998. Citado na página 24.

BORRA, C.; GONZÁLEZ, L.; SEVILLA, A. Birth timing and neonatal health. *American Economic Review*, v. 106, n. 5, p. 329–32, 2016. Citado na página 89.

BOU-HAMAD, I.; LAROCQUE, D.; BEN-AMEUR, H. A review of survival trees. *Statistics surveys*, Amer. Statist. Assoc., the Bernoulli Soc., the Inst. Math. Statist., and the . . . , v. 5, p. 44–71, 2011. Citado na página 29.

BRANCO, P.; TORGO, L.; RIBEIRO, R. P. A survey of predictive modeling on imbalanced domains. *ACM Computing Surveys (CSUR)*, ACM New York, NY, USA, v. 49, n. 2, p. 1–50, 2016. Citado na página 31.

BRASIL; SAÚDE, M. da. Atenção ao pré-natal de baixo risco. *Cadernos de Atenção Básica, n. 32*, Ministério da Saúde Brasília, DF, 2012. Citado 2 vezes nas páginas 52 e 54.

BREIMAN, L. Random forests. *Machine learning*, Springer, v. 45, n. 1, p. 5–32, 2001. Citado na página 28.

BREIMAN, L. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science*, Institute of Mathematical Statistics, v. 16, n. 3, p. 199–231, 2001. Citado na página 23.

BREIMAN, L. Using iterated bagging to debias regressions. *Machine Learning*, Springer, v. 45, n. 3, p. 261–277, 2001. Citado na página 28.

BUGELLI, A. et al. Health capabilities and the determinants of infant mortality in brazil, 2004–2015: an innovative methodological framework. *BMC public health*, Springer, v. 21, n. 1, p. 1–17, 2021. Citado na página 21.

CAMPBELL, F. et al. Early childhood investments substantially boost adult health. *Science*, American Association for the Advancement of Science, v. 343, n. 6178, p. 1478–1485, 2014. Citado na página 48.

CARD, D.; FENIZIA, A.; SILVER, D. *The health impacts of hospital delivery practices*. [S.l.], 2019. Citado 2 vezes nas páginas 89 e 97.

CARDOSO, R. C. A. et al. Infant mortality in a very low birth weight cohort from a public hospital in rio de janeiro, rj, brazil. *Revista Brasileira de Saúde Materno Infantil*, SciELO Brasil, v. 13, n. 3, p. 237–246, 2013. Citado 2 vezes nas páginas 15 e 45.

CARRILLO, B.; FERES, J. G. *Low Birth Weight and Infant Mortality: Lessons from Brazil* ▪ . [S.l.], 2017. Citado na página 48.

CASEY, B. M.; MCINTIRE, D. D.; LEVENO, K. J. The continuing value of the apgar score for the assessment of newborn infants. *New England Journal of Medicine*, Mass Medical Soc, v. 344, n. 7, p. 467–471, 2001. Citado na página 18.

CASTRO, M. C. et al. Brazil's unified health system: the first 30 years and prospects for the future. *The lancet*, Elsevier, v. 394, n. 10195, p. 345–356, 2019. Citado 2 vezes nas páginas 53 e 54.

CHAWLA, N. V.; JAPKOWICZ, N.; KOTCZ, A. Special issue on learning from imbalanced data sets. *ACM SIGKDD explorations newsletter*, ACM New York, NY, USA, v. 6, n. 1, p. 1–6, 2004. Citado na página 31.

CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. [S.l.: s.n.], 2016. p. 785–794. Citado 2 vezes nas páginas 29 e 30.

CHERKASSKY, V.; MULIER, F. M. *Learning from data: concepts, theory, and methods*. [S.l.]: John Wiley & Sons, 2007. Citado na página 25.

CHMIEL, F. et al. Using explainable machine learning to identify patients at risk of reattendance at discharge from emergency departments. *Scientific reports*, Nature Publishing Group, v. 11, n. 1, p. 1–11, 2021. Citado 2 vezes nas páginas 23 e 45.

CIESLAK, D. A.; CHAWLA, N. V. Learning decision trees for unbalanced data. In: SPRINGER. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. [S.l.], 2008. p. 241–256. Citado na página 31.

CONWAY, K. S.; DEB, P. Is prenatal care really ineffective? or, is the 'devil'in the distribution? *Journal of Health Economics*, Elsevier, v. 24, n. 3, p. 489–513, 2005. Citado 2 vezes nas páginas 49 e 67.

COOKSON, R. Evidence-based policy making in health care: what it is and what it isn't. *Journal of Health Services Research & Policy*, SAGE Publications Sage UK: London, England, v. 10, n. 2, p. 118–121, 2005. Citado na página 87.

COSTA-RAMÓN, A. M. et al. It's about time: Cesarean sections and neonatal health. *Journal of Health Economics*, Elsevier, v. 59, p. 46–59, 2018. Citado 2 vezes nas páginas 89 e 97.

COUTINHO, P. R. et al. Factors associated with low birth weight in a historical series of deliveries in campinas, brazil. *Revista da Associação Médica Brasileira*, SciELO Brasil, v. 55, p. 692–699, 2009. Citado 2 vezes nas páginas 49 e 68.

COX, D. R. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 34, n. 2, p. 187–202, 1972. Citado na página 24.

CURRIE, J. Inequality at birth: Some causes and consequences. *American Economic Review*, v. 101, n. 3, p. 1–22, 2011. Citado na página 48.

CURRIE, J.; GRUBER, J. Health insurance eligibility, utilization of medical care, and child health. *The Quarterly Journal of Economics*, MIT Press, v. 111, n. 2, p. 431–466, 1996. Citado na página 57.

DIAS, M. A. B. et al. The decision of women for cesarean birth: a case study in two units of the supplementary health care system of the state of rio de janeiro. *Ciencia & saude coletiva*, SciELO Brasil, v. 13, n. 5, p. 1521–1534, 2008. Citado na página 89.

DINIZ, C. S. G. et al. Por que as mulheres no setor privado têm gestações mais curtas no brasil? desvio à esquerda da idade gestacional, cesárea e inversão da disparidade esperada. *J Hum Growth Dev*, v. 26, n. 1, p. 33–40, 2016. Citado na página 94.

DIPRETE, T. A.; GANGL, M. 7. assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological methodology*, SAGE Publications Sage CA: Los Angeles, CA, v. 34, n. 1, p. 271–310, 2004. Citado na página 64.

DOMINGUES, R. M. S. M. et al. Process of decision-making regarding the mode of birth in brazil: from the initial preference of women to the final mode of birth. *Cadernos de saude publica*, SciELO Brasil, v. 30, p. S101–S116, 2014. Citado na página 89.

DRUMMOND, C.; HOLTE, R. C. et al. C4. 5, class imbalance, and cost sensitivity: why under-sampling beats over-sampling. In: CITESEER. *Workshop on learning from imbalanced datasets II*. [S.l.], 2003. v. 11, p. 1–8. Citado na página 31.

ESTABROOKS, A.; JO, T.; JAPKOWICZ, N. A multiple resampling method for learning from imbalanced data sets. *Computational intelligence*, Wiley Online Library, v. 20, n. 1, p. 18–36, 2004. Citado na página 31.

EVANS, W. N.; LIEN, D. S. The benefits of prenatal care: evidence from the pat bus strike. *Journal of Econometrics*, Elsevier, v. 125, n. 1-2, p. 207–239, 2005. Citado na página 49.

FALCÃO, I. R. et al. Factors associated with low birth weight at term: a population-based linkage study of the 100 million brazilian cohort. *BMC pregnancy and childbirth*, BioMed Central, v. 20, n. 1, p. 1–11, 2020. Citado 2 vezes nas páginas 49 e 61.

FAÚNDES, A.; CECATTI, J. G. Which policy for caesarian sections in brazil? an analysis of trends and consequences. *Health policy and planning*, Oxford University Press, v. 8, n. 1, p. 33–42, 1993. Citado na página 89.

FILHO, A. C.; BATISTA, A. F. D. M.; SANTOS, H. G. D. Data leakage in health outcomes prediction with machine learning. comment on "prediction of incident hypertension within the next year: Prospective study using statewide electronic health records and machine learning". *Journal of Medical Internet Research*, JMIR Publications Inc., Toronto, Canada, v. 23, n. 2, p. e10969, 2021. Citado na página 32.

FISCHER-RASMUSSEN, W.; TROLLE, D. Abdominal versus vaginal delivery in breech presentation. a retrospective study comparing 420 breech presentations and 9,291 cephalic presentations for infants weighing more than 2, 5000 g at birth. *Acta obstetricia et gynecologica Scandinavica*, v. 46, p. 69–76, 1967. Citado na página 87.

FONSECA, C. R. B. da et al. Adequacy of antenatal care and its relationship with low birth weight in botucatu, são paulo, brazil: a case-control study. *BMC pregnancy and childbirth*, BioMed Central, v. 14, n. 1, p. 1–12, 2014. Citado na página 49.

FRANÇA, G. V. et al. Coverage and equity in reproductive and maternal health interventions in brazil: impressive progress following the implementation of the unified health system. *International journal for equity in health*, Springer, v. 15, n. 1, p. 1–12, 2016. Citado na página 67.

FREITAS, F. et al. Late-onset sepsis and mortality among neonates in a brazilian intensive care unit: a cohort study and survival analysis. *Epidemiology & Infection*, Cambridge University Press, v. 147, 2019. Citado na página 15.

FRICK, K. D.; LANTZ, P. M. Selection bias in prenatal care utilization: An interdisciplinary framework and review of the literature. *Medical Care Research and Review*, Sage Publications Sage CA: Thousand Oaks, CA, v. 53, n. 4, p. 371–396, 1996. Citado na página 56.

FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. *The elements of statistical learning*. [S.l.]: Springer series in statistics New York, 2001. v. 1. Citado 2 vezes nas páginas 25 e 26.

GAMA, S. G. N. d. et al. Factors associated with caesarean section among primiparous adolescents in brazil, 2011-2012. *Cadernos de saude publica*, SciELO Brasil, v. 30, p. S117–S127, 2014. Citado na página 89.

GAMPER-RABINDRAN, S.; KHAN, S.; TIMMINS, C. The impact of piped water provision on infant mortality in brazil: A quantile panel data approach. *Journal of Development Economics*, Elsevier, v. 92, n. 2, p. 188–200, 2010. Citado na página 21.

GARCIA, L. P.; FERNANDES, C. M.; TRAEBERT, J. Risk factors for neonatal death in the capital city with the lowest infant mortality rate in brazil,. *Jornal de pediatria*, SciELO Brasil, v. 95, p. 194–200, 2019. Citado 3 vezes nas páginas 15, 22 e 45.

GAYNOR, M.; LAUDICELLA, M.; PROPPER, C. Can governments do it better? merger mania and hospital outcomes in the english nhs. *Journal of health economics*, Elsevier, v. 31, n. 3, p. 528–543, 2012. Citado na página 55.

GIJBELS, I. Censored data. *Wiley Interdisciplinary Reviews: Computational Statistics*, Wiley Online Library, v. 2, n. 2, p. 178–188, 2010. Citado na página 24.

GILPIN, L. H. et al. Explaining explanations: An overview of interpretability of machine learning. In: IEEE. *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)*. [S.l.], 2018. p. 80–89. Citado na página 45.

GLEZERMAN, M. Five years to the term breech trial: the rise and fall of a randomized controlled trial. *American journal of obstetrics and gynecology*, Elsevier, v. 194, n. 1, p. 20–25, 2006. Citado 2 vezes nas páginas 87 e 90.

GOLDANI, M. Z. et al. Trends in prenatal care use and low birthweight in southeast brazil. *American journal of public health*, American Public Health Association, v. 94, n. 8, p. 1366–1371, 2004. Citado 2 vezes nas páginas 49 e 68.

GONZAGA, I. C. A. et al. Prenatal care and risk factors associated with premature birth and low birth weight in the a capital in the brazilian northeast. *Ciencia & saude coletiva*, Associação Brasileira de Saúde Coletiva, v. 21, n. 6, 2016. Citado 2 vezes nas páginas 49 e 68.

GONZALEZ, F.; KUMAR, S. Prenatal care and birthweight in mexico. *Applied Economics*, Taylor & Francis, v. 50, n. 10, p. 1156–1170, 2018. Citado na página 50.

GRAMBSCH, P. M.; THERNEAU, T. M. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, Oxford University Press, v. 81, n. 3, p. 515–526, 1994. Citado na página 15.

GRANT, D. Physician financial incentives and cesarean delivery: new conclusions from the healthcare cost and utilization project. *Journal of health economics*, Elsevier, v. 28, n. 1, p. 244–250, 2009. Citado na página 95.

GROSSGLAUSER, M.; SANER, H. Data-driven healthcare: from patterns to actions. *European journal of preventive cardiology*, Sage Publications Sage UK: London, England, v. 21, n. 2_suppl, p. 14–17, 2014. Citado na página 14.

GROSSMAN, M. Front matter, the demand for health: a theoretical and empirical investigation. In: *The demand for health: a theoretical and empirical investigation*. [S.l.]: NBER, 1972. p. 20–0. Citado na página 61.

GRUBER, J.; KIM, J.; MAYZLIN, D. Physician fees and procedure intensity: the case of cesarean delivery. *Journal of health economics*, Elsevier, v. 18, n. 4, p. 473–490, 1999. Citado na página 95.

GUANAIS, F. C. The combined effects of the expansion of primary health care and conditional cash transfers on infant mortality in brazil, 1998–2010. *American Journal of Public Health*, American Public Health Association, v. 105, n. S4, p. S593–S599, 2015. Citado na página 20.

GUYON, I. et al. A scaling law for the validation-set training-set size ratio. *AT&T Bell Laboratories*, Citeseer, v. 1, n. 11, 1997. Citado na página 26.

HACKELING, G. *Mastering Machine Learning with scikit-learn*. [S.l.]: Packt Publishing Ltd, 2017. Citado na página 38.

HAHN, J.; HAM, J. C.; MOON, H. R. The hausman test and weak instruments. *Journal of Econometrics*, Elsevier, v. 160, n. 2, p. 289–299, 2011. Citado na página 56.

HANNAH, M. E. et al. Planned caesarean section versus planned vaginal birth for breech presentation at term: a randomised multicentre trial. *The Lancet*, Elsevier, v. 356, n. 9239, p. 1375–1383, 2000. Citado 3 vezes nas páginas 87, 90 e 97.

HEAGERTY, P. J.; ZHENG, Y. Survival model predictive accuracy and roc curves. *Biometrics*, Wiley Online Library, v. 61, n. 1, p. 92–105, 2005. Citado na página 30.

HILL, K. Frameworks for studying the determinants of child survival. *Bulletin of the World Health Organization*, SciELO Public Health, v. 81, p. 138–139, 2003. Citado na página 45.

HOO, Z. H.; CANDLISH, J.; TEARE, D. *What is an ROC curve?* [S.l.]: BMJ Publishing Group Ltd and the British Association for Accident . . . , 2017. Citado na página 30.

III, H. S. B. Physician demand for leisure: implications for cesarean section rates. *Journal of Health Economics*, Elsevier, v. 15, n. 2, p. 233–242, 1996. Citado na página 94.

ILIODROMITI, S. et al. Apgar score and the risk of cause-specific infant mortality: a population-based cohort study. *The Lancet*, Elsevier, v. 384, n. 9956, p. 1749–1755, 2014. Citado na página 97.

ISAIAH, A. et al. Weak instruments in iv regression: Theory and practice. *Annual Review of Economics*, 2018. Citado na página 58.

ISHWARAN, H. et al. Random survival forests. *The annals of applied statistics*, Institute of Mathematical Statistics, v. 2, n. 3, p. 841–860, 2008. Citado na página 29.

JABBAR, H.; KHAN, R. Z. Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study). *Computer Science, Communication and Instrumentation Devices*, p. 163–172, 2015. Citado na página 26.

JAMES, G. et al. *An introduction to statistical learning*. [S.l.]: Springer, 2013. v. 112. Citado na página 27.

JENSEN, V. M.; WÜST, M. Can caesarean section improve child and maternal health? the case of breech babies. *Journal of health economics*, Elsevier, v. 39, p. 289–302, 2015. Citado 4 vezes nas páginas 87, 88, 90 e 97.

JEWELL, R. T.; TRIUNFO, P. The impact of prenatal care on birthweight: the case of uruguay. *Health economics*, Wiley Online Library, v. 15, n. 11, p. 1245–1250, 2006. Citado 4 vezes nas páginas 49, 55, 56 e 66.

JURAJDA, Š. Lecture notes on identification strategies. 2007. Citado na página 67.

KAPLAN, E. L.; MEIER, P. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, Taylor & Francis, v. 53, n. 282, p. 457–481, 1958. Citado na página 24.

KLEINBERG, J. et al. Prediction policy problems. *American Economic Review*, v. 105, n. 5, p. 491–95, 2015. Citado na página 46.

KRAUSE, M. S.; HOWARD, K. I. What random assignment does and does not do. *Journal of Clinical Psychology*, Wiley Online Library, v. 59, n. 7, p. 751–766, 2003. Citado na página 93.

LABRECQUE, J.; SWANSON, S. A. Understanding the assumptions underlying instrumental variable analyses: a brief review of falsification strategies and related tools. *Current epidemiology reports*, Springer, v. 5, n. 3, p. 214–220, 2018. Citado na página 56.

LAWSON, G. W. The term breech trial ten years on: primum non nocere? *Birth*, Wiley Online Library, v. 39, n. 1, p. 3–9, 2012. Citado na página 87.

LEAL, M. d. C. et al. Prenatal care in the brazilian public health services. *Revista de saúde pública*, SciELO Brasil, v. 54, 2020. Citado 2 vezes nas páginas 54 e 55.

LEE, H. C.; SUBEH, M.; GOULD, J. B. Low apgar score and mortality in extremely preterm neonates born in the united states. *Acta Paediatrica*, Wiley Online Library, v. 99, n. 12, p. 1785–1789, 2010. Citado na página 97.

LEIBOWITZ, A. A. The demand for health and health concerns after 30 years. *Journal of Health Economics*, Elsevier, v. 23, n. 4, p. 663–671, 2004. Citado na página 67.

LIE, K. K.; GRØHOLT, E.-K.; ESKILD, A. Association of cerebral palsy with apgar score in low and normal birthweight infants: population based cohort study. *Bmj*, British Medical Journal Publishing Group, v. 341, 2010. Citado na página 97.

LO, J. C. Patients' attitudes vs. physicians' determination: implications for cesarean sections. *Social Science & Medicine*, Elsevier, v. 57, n. 1, p. 91–96, 2003. Citado na página 94.

LOPES, S. A. V. d. A. et al. Mortality for critical congenital heart diseases and associated risk factors in newborns. a cohort study. *Arquivos brasileiros de cardiologia*, SciELO Brasil, v. 111, p. 666–673, 2018. Citado na página 15.

LOTUFO, F. A. et al. Applying the new concept of maternal near-miss in an intensive care unit. *Clinics*, SciELO Brasil, v. 67, p. 225–230, 2012. Citado na página 97.

LUNDBERG, S. M.; LEE, S.-I. A unified approach to interpreting model predictions. In: *Proceedings of the 31st international conference on neural information processing systems*. [S.l.: s.n.], 2017. p. 4768–4777. Citado 2 vezes nas páginas 16 e 31.

MACHAREY, G. et al. Term cesarean breech delivery in the first pregnancy is associated with an increased risk for maternal and neonatal morbidity in the subsequent delivery: a national cohort study. *Archives of Gynecology and Obstetrics*, Springer, v. 302, n. 1, p. 85–91, 2020. Citado 2 vezes nas páginas 90 e 97.

MANI, I.; ZHANG, I. knn approach to unbalanced data distributions: a case study involving information extraction. In: ICML UNITED STATES. *Proceedings of workshop on learning from imbalanced datasets*. [S.l.], 2003. v. 126. Citado na página 31.

MCCALLUM, C. Explaining caesarean section in salvador da bahia, brazil. *Sociology of health & illness*, Wiley Online Library, v. 27, n. 2, p. 215–242, 2005. Citado na página 89.

MCGAUHEY, P. J. et al. Social environment and vulnerability of low birth weight children: a social-epidemiological perspective. *Pediatrics*, Am Acad Pediatrics, v. 88, n. 5, p. 943–953, 1991. Citado na página 61.

MEHRABI, N. et al. A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, ACM New York, NY, USA, v. 54, n. 6, p. 1–35, 2021. Citado na página 46.

MENDES, C. Q. d. S. et al. Low birth weight in a municipality in the southeast region of brazil. *Revista brasileira de enfermagem*, SciELO Brasil, v. 68, p. 1169–1175, 2015. Citado na página 49.

MENEZES, A. et al. Caesarean sections and risk of wheezing in childhood and adolescence: data from two birth cohort studies in brazil. *Clinical & Experimental Allergy*, Wiley Online Library, v. 41, n. 2, p. 218–223, 2011. Citado 2 vezes nas páginas 90 e 98.

MIAO, F. et al. Is random survival forest an alternative to cox proportional model on predicting cardiovascular disease? In: SPRINGER. *6TH European conference of the international federation for medical and biological engineering*. [S.l.], 2015. p. 740–743. Citado na página 15.

MIAO, F. et al. Predictive modeling of hospital mortality for patients with heart failure by using an improved random survival forest. *IEEE Access*, IEEE, v. 6, p. 7244–7253, 2018. Citado na página 15.

MODABBERNIA, A. et al. Apgar score and risk of autism. *European journal of epidemiology*, Springer, v. 34, n. 2, p. 105–114, 2019. Citado na página 97.

MONCADA-TORRES, A. et al. Explainable machine learning can outperform cox regression predictions and provide insights in breast cancer survival. *Scientific Reports*, Nature Publishing Group, v. 11, n. 1, p. 1–13, 2021. Citado 3 vezes nas páginas 15, 23 e 45.

MOSLEY, W. H.; CHEN, L. C. An analytical framework for the study of child survival in developing countries. *Population and development review*, JSTOR, v. 10, p. 25–45, 1984. Citado 5 vezes nas páginas 22, 23, 32, 33 e 44.

MÜHLRAD, H. Cesarean sections for high-risk births: health, fertility, and labor market outcomes. *The Scandinavian Journal of Economics*, Wiley Online Library, 2018. Citado 4 vezes nas páginas 87, 88, 90 e 97.

MUSAVI, M. T. et al. On the training of radial basis function classifiers. *Neural networks*, Elsevier, v. 5, n. 4, p. 595–603, 1992. Citado na página 39.

NAKAMURA-PEREIRA, M. et al. Use of robson classification to assess cesarean section rate in brazil: the role of source of payment for childbirth. *Reproductive health*, Springer, v. 13, n. 3, p. 245–256, 2016. Citado na página 94.

NASEJJE, J. B.; MWAMBI, H. Application of random survival forests in understanding the determinants of under-five child mortality in uganda in the presence of covariates that satisfy the proportional and non-proportional hazards assumption. *BMC research notes*, BioMed Central, v. 10, n. 1, p. 1–18, 2017. Citado 2 vezes nas páginas 15 e 34.

OBERMEYER, Z.; EMANUEL, E. J. Predicting the future—big data, machine learning, and clinical medicine. *The New England journal of medicine*, NIH Public Access, v. 375, n. 13, p. 1216, 2016. Citado na página 25.

OGAWA, K. et al. Association between adolescent pregnancy and adverse birth outcomes, a multicenter cross sectional japanese study. *Scientific Reports*, Nature Publishing Group, v. 9, n. 1, p. 1–8, 2019. Citado na página 43.

O'NEILL, T. J. Inconsistency of the misspecified proportional hazards model. *Statistics & probability letters*, Elsevier, v. 4, n. 5, p. 219–222, 1986. Citado na página 24.

ORGANIZATION, W. H. et al. *WHO recommendations on antenatal care for a positive pregnancy experience*. [S.l.]: World Health Organization, 2016. Citado na página 48.

PAIM, J. et al. The brazilian health system: history, advances, and challenges. *The Lancet*, Elsevier, v. 377, n. 9779, p. 1778–1797, 2011. Citado 2 vezes nas páginas 20 e 54.

PAIXAO, E. S. et al. Risk of mortality for small newborns in brazil, 2011-2018: A national birth cohort study of 17.6 million records from routine register-based linked data. *The Lancet Regional Health-Americas*, Elsevier, v. 3, p. 100045, 2021. Citado na página 45.

PAIXAO, E. S. et al. Associations between cesarean delivery and child mortality: A national record linkage longitudinal study of 17.8 million births in brazil. *PLoS medicine*, Public Library of Science San Francisco, CA USA, v. 18, n. 10, p. e1003791, 2021. Citado 4 vezes nas páginas 15, 33, 90 e 96.

PINHEIRO, C. E. A.; PERES, M. A.; D'ORSI, E. Increased survival among lower-birthweight children in southern brazil. *Revista de saude publica*, SciELO Brasil, v. 44, p. 776–784, 2010. Citado 2 vezes nas páginas 15 e 45.

PÖLSTERL, S.; NAVAB, N.; KATOUZIAN, A. Fast training of support vector machines for survival analysis. In: SPRINGER. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. [S.l.], 2015. p. 243–259. Citado na página 27.

PROBST, P.; BOULESTEIX, A.-L.; BISCHL, B. Tunability: importance of hyperparameters of machine learning algorithms. *The Journal of Machine Learning Research*, JMLR. org, v. 20, n. 1, p. 1934–1965, 2019. Citado na página 26.

RAIFMAN, S.; CUNHA, A. J.; CASTRO, M. C. Factors associated with high rates of caesarean section in brazil between 1991 and 2006. *Acta Paediatrica*, Wiley Online Library, v. 103, n. 7, p. e295–e299, 2014. Citado na página 88.

RAMOS, R. et al. Using predictive classifiers to prevent infant mortality in the brazilian northeast. In: IEEE. *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*. [S.l.], 2017. p. 1–6. Citado na página 15.

RANDOLPH, J. J.; FALBE, K. A step-by-step guide to propensity score matching in r. *Practical Assessment, Research & Evaluation*, Citeseer, v. 19, 2014. Citado na página 64.

REFAEILZADEH, P.; TANG, L.; LIU, H. Cross-validation. *Encyclopedia of database systems*, Springer, v. 5, p. 532–538, 2009. Citado na página 26.

RIETBERG, C. C. T.; ELFERINK-STINKENS, P. M.; VISSER, G. H. The effect of the term breech trial on medical intervention behaviour and neonatal outcome in the netherlands: an analysis of 35,453 term breech infants. *BJOG: An International Journal of Obstetrics & Gynaecology*, Wiley Online Library, v. 112, n. 2, p. 205–209, 2005. Citado 2 vezes nas páginas 87 e 90.

RISSO, S. d. P.; NASCIMENTO, L. F. C. Risk factors for neonatal death in neonatal intensive care unit according to survival analysis. *Revista Brasileira de terapia intensiva*, SciELO Brasil, v. 22, p. 19–26, 2010. Citado 2 vezes nas páginas 15 e 45.

ROGERS, J.; GUNN, S. Identifying feature relevance using a random forest. In: SPRINGER. *International Statistical and Optimization Perspectives Workshop"Subspace, Latent Structure and Feature Selection"*. [S.l.], 2005. p. 173–184. Citado na página 30.

ROSENBAUM, P. R.; RUBIN, D. B. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 45, n. 2, p. 212–218, 1983. Citado na página 65.

ROYSTON, P.; PARMAR, M. K. An approach to trial design and analysis in the era of non-proportional hazards of the treatment effect. *Trials*, Springer, v. 15, n. 1, p. 1–10, 2014. Citado na página 15.

RUSSO, L. X. et al. Primary care physicians and infant mortality: evidence from brazil. *PLoS One*, Public Library of Science San Francisco, CA USA, v. 14, n. 5, p. e0217614, 2019. Citado na página 20.

SANTOS, G. R. B. d. et al. Relationship between microcephaly and zika virus during pregnancy: a review. *Revista da Associação Médica Brasileira*, SciELO Brasil, v. 64, p. 635–642, 2018. Citado na página 62.

SCHAPIRE, R. E. A brief introduction to boosting. In: CITESEER. *Ijcai*. [S.l.], 1999. v. 99, p. 1401–1406. Citado 2 vezes nas páginas 28 e 29.

SCHEMPER, M. Cox analysis of survival data with non-proportional hazard functions. *Journal of the Royal Statistical Society: Series D (The Statistician)*, Wiley Online Library, v. 41, n. 4, p. 455–465, 1992. Citado na página 37.

SCHOENFELD, D. Partial residuals for the proportional hazards regression model. *Biometrika*, Oxford University Press, v. 69, n. 1, p. 239–241, 1982. Citado na página 25.

SCHULKIND, L.; SHAPIRO, T. M. What a difference a day makes: quantifying the effects of birth timing manipulation on infant health. *Journal of health economics*, Elsevier, v. 33, p. 139–158, 2014. Citado na página 90.

SEMYKINA, A.; WOOLDRIDGE, J. M. Estimating panel data models in the presence of endogeneity and selection. *Journal of Econometrics*, Elsevier, v. 157, n. 2, p. 375–380, 2010. Citado na página 57.

SHAPLEY, L. S.; ROTH, A. E. et al. *The Shapley value: essays in honor of Lloyd S. Shapley*. [S.l.]: Cambridge University Press, 1988. Citado na página 30.

SILVEIRA, D. S.; SANTOS, I. S. Adequacy of prenatal care and birthweight: a systematic review. *Cadernos de saude publica*, SciELO Brasil, v. 20, n. 5, p. 1160–1168, 2004. Citado na página 49.

SIMÕES, R. et al. Elective cesarean section for term breech delivery. *Revista da Associação Médica Brasileira*, SciELO Brasil, v. 61, p. 391–402, 2015. Citado 3 vezes nas páginas 91, 96 e 97.

SOMMER, J.; SARIGIANNIS, D.; PARNELL, T. Learning to tune xgboost with xgboost. *arXiv preprint arXiv:1909.07218*, 2019. Citado na página 39.

SONG, Y.-Y.; YING, L. Decision tree methods: applications for classification and prediction. *Shanghai archives of psychiatry*, Shanghai Mental Health Center, v. 27, n. 2, p. 130, 2015. Citado na página 28.

SOUZA, J. P. et al. Caesarean section without medical indications is associated with an increased risk of adverse short-term maternal outcomes: the 2004-2008 who global survey on maternal and perinatal health. *BMC medicine*, BioMed Central, v. 8, n. 1, p. 1–10, 2010. Citado na página 98.

SOUZA, S. de; DUIM, E.; NAMPO, F. K. Determinants of neonatal mortality in the largest international border of brazil: a case-control study. *BMC Public Health*, BioMed Central, v. 19, n. 1, p. 1–9, 2019. Citado na página 23.

SPINOLA, P. d. S. L. Convenience effect on birth timing manipulation: Evidence from brazil. *Programa de Pós-Graduação em Economia, PPGE/IE-UFRJ*, 2016. Citado na página 94.

STUART, E. A. Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, NIH Public Access, v. 25, n. 1, p. 1, 2010. Citado 2 vezes nas páginas 64 e 94.

STUART, E. A.; LEE, B. K.; LEACY, F. P. Prognostic score–based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. *Journal of clinical epidemiology*, Elsevier, v. 66, n. 8, p. S84–S90, 2013. Citado na página 65.

SZWARCWALD, C. L. et al. Inequalities in infant mortality in brazil at subnational levels in brazil, 1990 to 2015. *Population Health Metrics*, Springer, v. 18, n. 1, p. 1–9, 2020. Citado na página 21.

TIAN, Y.; ZHANG, Y. A comprehensive survey on regularization strategies in machine learning. *Information Fusion*, Elsevier, v. 80, p. 146–166, 2022. Citado na página 39.

TOIJONEN, A. et al. Neonatal outcome in vaginal breech labor at 32+ 0—36+ 0 weeks of gestation: a nationwide, population-based record linkage study. *BMC Pregnancy and Childbirth*, BioMed Central, v. 22, n. 1, p. 1–7, 2022. Citado na página 91.

TOMASI, E. et al. Qualidade da atenção pré-natal na rede básica de saúde do brasil: indicadores e desigualdades sociais. *Cadernos de saúde pública*, SciELO Brasil, v. 33, 2017. Citado na página 55.

TRESP, V. Committee machines. *Handbook for neural network signal processing*, CRC Press Boca Raton, FL, p. 1–18, 2001. Citado na página 27.

TURNER, M. The term breech trial: Are the clinical guidelines justified by the evidence? *Journal of obstetrics and gynaecology*, Taylor & Francis, v. 26, n. 6, p. 491–494, 2006. Citado na página 90.

VALENTE, M. P.; AFONSO, M. C.; CLODE, N. Is vaginal breech delivery still a safe option? *Revista Brasileira de Ginecologia e Obstetrícia*, SciELO Brasil, v. 42, p. 712–716, 2020. Citado 2 vezes nas páginas 91 e 97.

VALTER, R. et al. Data mining and risk analysis supporting decision in brazilian public health systems. In: IEEE. *2019 IEEE International Conference on E-health Networking, Application & Services (HealthCom)*. [S.l.], 2019. p. 1–6. Citado na página 15.

VELOSO, H. J. F. et al. Low birth weight in são luís, northeastern brazil: trends and associated factors. *BMC pregnancy and childbirth*, BioMed Central, v. 14, n. 1, p. 1–12, 2014. Citado 2 vezes nas páginas 49 e 68.

VICTORA, C. G. et al. Maternal and child health in brazil: progress and challenges. *The lancet*, Elsevier, v. 377, n. 9780, p. 1863–1876, 2011. Citado na página 67.

VIELLAS, E. F. et al. Assistência pré-natal no brasil. *Cadernos de Saúde Pública*, SciELO Brasil, v. 30, p. S85–S100, 2014. Citado 2 vezes nas páginas 20 e 55.

WANG, B.-s. et al. Effects of caesarean section on maternal health in low risk nulliparous women: a prospective matched cohort study in shanghai, china. *BMC pregnancy and childbirth*, BioMed Central, v. 10, n. 1, p. 1–10, 2010. Citado na página 98.

WANG, W.; TEMSAH, G.; MALLICK, L. The impact of health insurance on maternal health care utilization: evidence from ghana, indonesia and rwanda. *Health policy and planning*, Oxford University Press, v. 32, n. 3, p. 366–375, 2017. Citado 2 vezes nas páginas 56 e 57.

WEHBY, G. L. et al. Prenatal care effectiveness and utilization in brazil. *Health policy and planning*, Oxford University Press, v. 24, n. 3, p. 175–188, 2009. Citado 5 vezes nas páginas 49, 55, 56, 66 e 67.

WEHBY, G. L.; ULLRICH, F.; XIE, Y. Very low birth weight hospital volume and mortality–an instrumental variables approach. *Medical care*, NIH Public Access, v. 50, n. 8, p. 714, 2012. Citado na página 57.

WEI, L.-J. The accelerated failure time model: a useful alternative to the cox regression model in survival analysis. *Statistics in medicine*, Wiley Online Library, v. 11, n. 14-15, p. 1871–1879, 1992. Citado na página 29.

YUSSIF, A.-S. et al. The long-term effects of adolescent pregnancies in a community in northern ghana on subsequent pregnancies and births of the young mothers. *Reproductive health*, Springer, v. 14, n. 1, p. 1–7, 2017. Citado na página 43.

ZEILEIS, A. Econometric computing with hc and hac covariance matrix estimators. Institut für Statistik und Mathematik, WU Vienna University of Economics and . . . , 2004. Citado na página 58.

ZIMMERMAN, E.; WOOLF, S. H. Understanding the relationship between education and health. *NAM Perspectives*, 2014. Citado na página 61.

APÊNDICES

ANEXOS