ABSTRACT

Title of Dissertation: ESSAYS ON THE ECONOMICS OF CRIME, GENDER, AND HEALTH

Elena Ramirez Pierce
Doctor of Philosophy, 2023

Dissertation Directed by: Professor Jessica Goldberg
Department of Economics

In this dissertation I study the impacts of large government programs on crime and health outcomes. I also run an online experiment about the use of professional titles to elicit perceptions of experts cited in major news outlets and to test whether these perceptions vary across genders.

In the first chapter, I examine the effects of large changes in cash availability on crime. South Africa has a large social safety net comprised of numerous cash transfer programs, called social grants, that are paid on a monthly basis. Prior to late 2018, these social grants were paid mostly in cash at grant disbursement locations called paypoints. Using a differences-in-differences (DiD) strategy, I analyze the effects of the temporary increase in cash availability on crime by comparing crimes on social grant payment dates in small geographical areas, police precincts, between areas with differing numbers of cash disbursement locations. The results suggest a small decrease in crime the day prior to social grant payments, and small increases the day of payments or the day after payments, depending on the empirical specification. These results are consistent with perpetrators potentially delaying their labor supply of crime until the widely publicized cash

grant payment days, an anticipation effect, and increasing their labor supply of crime on or after payment days consistent with a loot effect, resulting from increased cash and purchased goods availability.

Chapter two investigates whether there exists a credibility penalty for female experts compared to male experts when major news outlets forgo the use of professional titles, such as “Dr.” that serve as an information signal on the level of their training. Given the extensive literature on gender and racial bias in media reporting and professional and academic environments, the practice of abstaining from the use of professional titles may reinforce and even exacerbate these biases. In this co-authored analysis, we test for differential effects by conducting an online experiment that presented survey respondents with news articles holding constant content, but varying the gender and title of the cited experts and asked them to rate the expert’s credibility. Our design enables between-subject and within-subject analysis. While we are able to detect a positive credibility effect of using professional titles, we are unable to distinguish a differential credibility impact across gender.

Finally, in chapter three I estimate the effects of a large-scale national physician provision program in Brazil on birth outcomes. Given the risk to mothers of injury and disease associated with childbirth that may affect the health of the newborn, as well as the myriad of complications that may arise that could threaten the health of the fetus, increasing access to and quality of medical care may have substantial effects on birth outcomes. The Mais Medicos Program (PMM) focused on equalizing physicians per capita as well as generally increasing the number of physicians across the country. Beginning in late 2013 and an executive branch initiative, the program placed almost 20,000 physicians by 2016, predominantly from Cuba, throughout the country. Using vital statistics data of the universe of births in Brazil from 2006 to 2017, I estimate the effect of increasing the supply of primary care physicians on birth weight using both a differences-in-differences and an instrumental variables approach. I find that PMM resulted in higher average birth weight for children throughout Brazil. However, I find no improvement on the incidence of low birth weight or any weight effects for those living in rural parts of the country. Hence, these results imply PMM did not affect the most vulnerable pregnancies.

ESSAYS ON THE ECONOMICS OF CRIME, GENDER, AND HEALTH

by

Elena Ramirez

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Associate Professor Jessica Goldberg, Chair/Advisor
Professor Sebastian Galiani, Co-Advisor
Associate Professor Ethan Kaplan
Assistant Professor Nolan Pope
Professor Susan Parker

Foreword

Chapter 3 of this dissertation, titled "The "Dr." Effect", includes jointly authored work with Zia Saylor of Williams College, and Mary Yilma of the Massachusetts Institute of Technology. As the principal investigator of this work, I, Elena Ramirez, as approved by my dissertation committee, have made substantial contributions to the relevant aspects of this chapter, including survey design and implementation, data collection, empirical analysis, and writing.

Dedication

To my parents Mario Ramirez Pallares and Maria Elena Ramirez Ortiz

Acknowledgments

I am incredibly grateful for all of the guidance and support my advisor Jessica Goldberg has provided me over the last 10 years. She has provided invaluable advice throughout my undergraduate and graduate studies, as well as helped me navigate life outside of school. I am thankful to her for always being in my corner and for being my role-model all of these years. I am also grateful to her for allowing me the opportunity to serve as the graduate assistant for the Promoting Achievement and Diversity in Economics (PADE) program for several years, giving me the opportunity to mentor and guide undergraduate students. Developing relationships with talented and ambitious undergraduate students was extremely rewarding and I will value those relationships for years to come.

I am grateful to several other faculty and staff members in the Economics department. I thank my co-advisor Sebastian Galiani for providing vital feedback throughout my doctoral studies. I also want to thank Ethan Kaplan for taking the time to be on my committee, but also for showing me what it means to be a kind, fair, and impactful teacher as I had the opportunity to be his teaching assistant for numerous courses. I would also like to thank Nolan Pope and Susan Parker for contributing their time and expertise to my dissertation committee. I am also grateful to Melissa Kearney for providing guidance and feedback, particularly in the early stages of my research, and for also being a role-model to so many of us in the department. I also want to thank Guido Kuersteiner for pushing me and providing support during a difficult time. Thank you to Vickie Fletcher for helping us overcome everything from administrative to personal obstacles. I also want to thank Susan Godlonton at Williams College for providing feedback and expertise on both my experimental research as well as my research in South Africa. Thank you to Mary Yilma and Zia Saylor for all of their time and effort into our experimental research, and for teaching me many lessons along the way.

I am extremely grateful for the friendships I have gained throughout the doctoral program. In particular, without Lea Rendell, Roberto Lagos, and Victoria Perez-Zetune, I would not have survived the first year. The countless hours we spent supporting and motivating each other are some of my most treasured memories. I also want to thank my friends Chris Roudiez, Mrin Chaterjee, Luke Pardue, Nathalie Gonzalez Prieto, Shuqing Chen, Macarena Kutscher, and Anusuya Sivaram for all of the research advice, life advice, and fun memories we've shared throughout these years.

I cannot say thank you enough to my support system, my parents, my siblings Mario, Jorge Luis, and Michelle, my dearest cousin Stephanie Ortiz, my niece and nephews Abby, David, and Junior, my friends Camille, Lydia, Anna, Georgia, Rachel, and Aaron. Thank you for being with me through the ups and downs these past years, without your constant love and support I would not be who I am today. I also want to thank my husband Robbie. Thank you for being a part of this journey from start to finish, and pushing me, motivating me, comforting me, and loving me unconditionally throughout.

Table of Contents

Bibliography 151

List of Tables

List of Figures

List of Abbreviations

AP	Associated Press
CAS	Crime Administrative System
CCT	Conditional Cash Transfer
CPS	Cash Payment Services
DiD	Differences-in-Differences
Dr	Doctor
FBI	Federal Bureau of Investigation
GDP	Gross Domestic Product
IPW	Inverse Probability Weighting
ISS	Institute for Security Studies
LGBTQ+	Lesbian, Gay, Bisexual, Transgender, Queer, and Other
NOAA	National Oceanic and Atmospheric Administration
PhD	Doctor of Philosophy
PIH	Permanent Income Hypothesis
PIN	Personal Identification Number
PMM	Programma Mais Medicos (More Doctors Program)
RCT	Randomized Control Trial
SAPO	South African Post Office
SAPS	South African Police Service
SASSA	South African Social Security Agency
UCR	Uniform Crime Reporting
US	United States
VoCS	Victims of Crime Survey

Chapter 1: Cash and Crime: Evidence from Large Government Cash Transfers

1.1 Introduction

South Africa has large social protection programs that prior to late 2018 made regularly scheduled payments in cash at particular locations, called paypoints, on well-publicized dates at the beginning of each month. This payment method varied substantially from other large government cash transfer programs such as Bolsa Familia in Brazil or Prospera in Mexico that at the same time made payments via debit cards or direct deposits into bank accounts, forgoing cash altogether. The cash transfer programs provide grants of significant value to over a third of South Africans today. South Africa is also plagued with high crime rates, in the extreme case, suffering from a murder rate of 36.2 per 100,000 people in 2019, one of the highest in the world. The features of the program potentially created an environment that increased crime. By creating temporarily high returns to criminal activity through periodic spikes in cash availability, the program's payment mechanism may inadvertently increase crime. This mechanism, operating through the expected return to criminal activity, is consistent with the Becker seminal model of crime [1].

Through the economic framework based on increased expected returns to crime, I exploit the variation in the timing and location of both the cash grant payments and criminal incidents to study how temporary changes to cash availability affected local crime. I focus on the South African social grant programs that made payments in cash at specific physical disbursement locations on or around the first day of each month. This method of payments resulted in not only an increase in the amount of cash available, but also an increase in goods purchased with the cash available for theft, termed the "loot" effect in the literature [2, 3]. Multiple media reports highlight the occurrence of crime around paypoint locations, many detailing reports of armed robberies, assaults, and murders on social grant payment days [4–9].

The social grants represent a substantial portion of government expenditure and beneficiaries' household budgets. In 2019, the programs accounted for 3.2 percent of South African Gross Domestic Product (GDP). Findings from a FinScope survey in 2018 indicate that 58 percent of child and foster child support grant recipients, 93 percent of old age grant recipients, and 89 percent of disability grant recipients consider the grant as their main source of income, respectively [10]. Moreover, the World Bank estimates that for the poorest 20 percent of households, the grants represent about 60 percent of their household expenditure [11]. Therefore, the payment mechanism of the social grant programs prior to 2018 provided a unique setting to study the relationship between cash and crime.

To estimate the causal impact of the change in the expected returns to crime on crime levels, I employ differences-in-differences (DiD) specifications that exploit the temporal and geographic characteristics of crimes and cash disbursements sites and study how the number of physical cash disbursement sites affect local crime levels on and around social grant payment days. Comparing small geographic areas (police precincts) with and without paypoint locations and leveraging the timing of benefit distribution, I am able to estimate the effect of temporarily increasing cash availability, and therefore increasing returns to crime, on local crime levels. Given the source of variation I exploit in this analysis, I am unable to disentangle between the mechanisms affecting potential offenders and social grant recipients but instead estimate a combined reduced-form effect.

I test alternative specifications exploiting the variation in the number of cash disbursement sites across police precincts by comparing places with above and below median number of paypoint locations as well as by using the number of paypoints across police precincts to estimate the effect of one additional paypoint location.

Across specifications, I find small effects of a decrease in crime on the day prior to social grant payments and an increase in crime on or the day after the payments. When I define treatment as either police precincts with paypoint locations or precincts with above median paypoint locations, though imprecisely estimated, I observe a decrease in average daily total crime of between 5 and 7 percent from the control group mean. I also observe an increase in crime the day after payments when comparing precincts with and without paypoints, and an increase on the day of payments when comparing precincts with above and below the median number of paypoints. Additionally, I estimate a generalized DiD specification but find only quantitatively small effects of an additional paypoint location on crime on or around social grant payment days. These results are robust to numerous sample restrictions determining that particular types of places do not drive the results. Finally, results from a falsification exercise using monthly salary paydays instead of social grant paydays do not replicate the anticipation effects observed from social grant payments, but have similar magnitude increases in crime on and after grant payments. This suggest that cash, whether from grants or other sources, may have a different effect in places where grant paypoints are located.

1.2 Relationship Between Cash and Crime

Changes to cash availability around particular locations in communities, through government transfer programs or other means, can affect crime via several mechanisms. These include through changes to the expected returns to financially motivated crimes and therefore the behavior of potential perpetrators, and through changes to the liquidity of beneficiaries that may affect their behavior in ways that change their likelihood of victimization and engagement with crime. Economists and criminologists argue that cash has a unique relationship with crime, such that its use as the modality of payment for government transfer programs has effects on crime different to other payment methods such as Electronic Benefit Transfers (EBT) or direct deposits into personal bank accounts. Firstly, cash as a form of payment may directly affect returns to crime as it is a desirable good to perpetrators due to its liquidity and ease of anonymous expenditure, thereby reducing the likelihood of their apprehension [12–15]. Additionally, cash has durable value, such that it may be stored for future anonymous consumption, unlike stolen credit cards, for example, that can be cancelled following theft and may incriminate the offender if used [16]. Finally, cash may also represent the only means of exchange in certain criminal environments, such as underground economies, for the consumption of illicit goods and services [17–19]. Ultimately, cash is highly and uniquely tempting to criminals for the reasons outlined above, and therefore cash as the form of payment for government transfers may have a significant impact on crime.

Regularly scheduled cash transfers at physical locations increases the return to crime and may consequently change potential offender's behavior on those payment days around paypoint locations. Most similar to the current analysis, Gandelman, Munyo and Schertz also exploit temporal and geographic variation in cash availability and find that the prohibition of cash payments at gas stations in Uruguay during certain hours of the night resulted in large decreases in the rate of robberies, between 30 and 50 percent [20]. This work highlights the potential loot effect created by cash, with no scope for effects due to potential victims' behavior. In a setting similar to the South Africa social grant programs, Borraz and Munyo analyze the effect of increasing cash availability through a restructuring of a conditional cash transfer (CCT) program in Uruguay in 2008, exploiting only the timing of the change in payment mechanism, and find that transfer payments in cash lead to an increase in crime, particularly for property crimes, and therefore financially motivated crimes against beneficiaries [3]. Again, this work addresses the potential for a loot effect from cash, here combining the effects on both perpetrators and victims of crime. Wright et al., analyzing the impact of cash availability on the same outcomes as in my analysis - total crimes and violent crimes - exploit only variation in timing and find that the switch from check payments to EBT payments in Missouri led to decreases across these outcomes [21]. Pridemore, Roche, and Rogers also study decreased cash availability through non-cash government transfers. They use a cross-country analysis, and find that lower rates of cash disbursement are associated with lower incidence of financially motivated crimes [16]. Finally, in another cross-country analysis, Armey, Lipow, and Webb find that the global shift away from cash and towards electronic payments through increased use of credit and debit payments at point of sale is associated with a significant reduction in financially motivated crimes such as robberies and burglaries [22]. Their analysis also exploits only temporal variation. All of these analyses estimate a negative effect of decreasing cash availability on crime, though only one introduces geographical variation. Their results suggest that increasing cash availability around paypoint locations in South Africa should increase crimes, particularly financially motivated crimes, in those areas.

The second potential mechanism through which temporarily changing cash availability can affect crime is through changes in the liquidity of beneficiaries and therefore their behavior around payment dates and around paypoint locations. Illiquidity right before payments may drive them to commit financially motivated crimes. However, I would only observe this effect in my analysis if these crimes occurred around paypoint locations. This mechanism is explored in work by Carr and Packham using only variation in the timing of transfer payments. They find that changing from beginning-of-month SNAP benefit payments to a staggered schedule designed to alleviate beneficiaries' end-of-the-month resource constraints results in large reductions in crime, particularly thefts, at grocery stores [23]. However, I am unable to observe crimes such as shoplifting in my data, and therefore cannot provide evidence for the effects of this particular mechanism on crime.

Another means by which changing cash availability can affect crime through beneficiary behavior is if beneficiaries use their cash payments to increase their consumption of complementary goods to crime, such as drugs and alcohol that could make them more vulnerable to crime or even become perpetrators themselves. Criminologists outline a framework through which cash affects crime through the need of cash for consumption of drugs and alcohol [24–26]. Additionally, studies across a wide range of disciplines, sociology, psychology, and economics, among others, have documented the link between alcohol consumption and increases in violent crime [27–31]. In an examination of the effects of a yearly UBI payment from the Alaskan permanent fund, Watson, Guettabi and Reimer, exploiting only temporal variation, find a 14 percent increase in drug and alcohol related incidents across the state that persists for about a month after payments, suggesting that beneficiaries may use their transfer benefits for the consumption of what is described in the literature as temptation goods [32]. Dobkin and Puller find similar effects of increased consumption of drugs and alcohol using the temporal variation from monthly receipt of welfare payments in the US [33]. Castellari et al. use scanner data on a panel of household purchases and document increased purchases of beer in the US when SNAP benefits are paid closer to the weekend, even though the benefits cannot themselves be used for alcohol [34]. While none of these analyses focus on cash as the modality of payment, they still highlight how temporary changes to transfer recipients' liquidity can affect their consumption of goods complementary to crime. One limitation of the current analysis is that crimes related to drug and alcohol consumption are excluded from my data. Also, I could only detect crime due to this mechanism if beneficiaries engage in criminal behavior related to these goods near paypoint locations. While plausible, I only observe these effects if drug and alcohol activity result in violent crimes that are recorded as assaults, due to the nature of the data.

This analysis contributes to the literature by utilizing variation not only in the timing of cash transfer payments, as most of the literature previously described, but also in the precise geographical location of the payments, allowing for improved identification of the effect of cash on crime. Also, while much of the literature on cash transfers and crime has focused on how income effects from transfer programs affect beneficiaries' behaviors as offenders, like work by Tuttle (2019) finding increased rates of recidivism due to the loss of SNAP benefits [35], this potential mechanism is not relevant in this setting since I examine temporary changes to cash availability, not income, in particular geographical areas.

Moreover, I provide evidence on the relationship between cash and crime in a context where it may be especially important. I study a cash transfer program that is larger in value than most in the literature and takes place in a setting with high underlying crime levels. These features of the South African context may result in the potential for large observed effects. Thus, the estimated small, though imprecisely estimated, results on crime are unexpected and in contrast to the existing literature. These results could be due to the fact that most social grant beneficiaries are of old age or children, meaning they may have different propensities to be either victims or perpetrators of crimes compared to recipients of other programs in the existing literature.

1.3 Context

South Africa is a country with a large, robust, and long-established social welfare system. It has a large safety net comprised of several social grants paid on a monthly basis: old age pensions, disability grants, child support grants, care dependency grants, foster care grants, and war veterans' grants. The large safety net was created at the end of apartheid and designed to target poverty alleviation and reduce racial inequality. As of 2020, 18.5 million South Africans received social grants, from a population of 58.8 million, meaning almost a third of South Africans received a social grant. The social safety net is administered by the South African Social Security Agency (SASSA). The largest grant in terms of value to the recipient is the old age pension, for persons aged 60 and older, and is means-tested. The largest grant in terms of coverage is the child support grant, designed for children under the age of 18 and given to the primary caregiver, and is also means-tested. The third largest grant is the disability grant, for adults with medically certified disabilities. The trend in stock of social grant recipients is shown in Figure 1.2 using data from yearly SASSA reports on number of recipients and value of grants. The value and expenditures for each of these grants in shown in Table 1.1. The two largest grants, the old age pension and the child support grant have similar expenditures, at 83.1 and 84.9 billion South African Rands, though the child support grant has almost four times the number of beneficiaries in 2020.

SASSA paid social grants to recipients in one of three ways: cash at a paypoint station on the first of the month using biometric identification[1]; direct deposit into personal bank accounts; and SASSA gold cards, similar to EBT cards in the US. Prior to the mass decommissioning of paypoints at the end of 2018, recipients by far preferred collecting their grants in person in cash at paypoints [36]. South Africa, still, is considered a cash driven economy with over half of all consumer transactions paid in cash and with only 30 percent of bank accounts used more than three times in 30 days [10, 37, 38]. Grant recipients reported preferring cash because they faced high banking fees and problematic automatic deductions from banking institutions. Media reports, investigations by Black Sash, a partner of the South African Association for Community Advice Offices, and speeches in parliament by past Ministers of Social Development detail the seemingly widespread practice of banking institutions such as Grindoid bank making fraudulent deductions for airtime on mobile phones, funeral insurance policies, and other unexplained sums from the debit cards provided to SASSA beneficiaries for grant receipt [39–41]. Additionally, recipients face fees from banks to change their account PINs, see their account balances, obtain account statements, and withdraw funds. Following a move by South African banks to increase fees in 2015, general uproar prompted politicians to call for "drastic action" for banks to reduce their fees, while a 2016 report by Boston Consulting Group estimated South African bank fees to be four times higher than other developed markets, and emerging markets such as India [42, 43]. One of the five major banks, First National Bank, for example, charged about 3.2 percent of the value of the child support grant in 2014 to deposit it into an account, and similar charges for any withdrawals of the grant amount [44]. Additionally, mobile money products have low take-up rates compared to other African countries. In 2010, the mobile money app M-Pesa commenced operation in South Africa with a goal of subscribing 10 million users within three years [37].

[1]Biometric identification has been in use in South Africa since 2012.

However, several years later, only 76,000 users were registered, a markedly different experience than the success observed in Kenya [37].

Paypoint locations where recipients could collect cash on designated days were located throughout the country, in places such as town halls, convenience stores, or mobile pop-ups in towns and villages. Paypoints were concentrated in urban areas with higher population densities, and recipients in rural areas had to travel longer distances to obtain their grants, as depicted in Figure 1.5. The geographical distribution of paypoint locations and police stations across police precincts also shows that in the more rural and mountainous western areas, paypoints were located very close to police stations. This pattern is not as clear in the urban areas of the east. Unsatisfactory conditions for recipients at paypoint locations have long been documented by the media. Many news reports describe long lines at paypoints that sometimes commence the night before, or early the morning of the scheduled grant payment date [36, 45, 46]. This flocking of millions of recipients to designated cash disbursement sites throughout the country could have increased their likelihood of victimization by clearly signaling individuals' upcoming receipt of large cash sums.

South Africa also suffers from high crime rates, including violent crimes. Extremely violent crimes, such as murder, occur at lower levels than non-violent crimes but are likely the best recorded crimes. South Africa experienced a murder rate of 36.2 per 100,000 persons in 2019. In the same year, Brazil and the US experienced murder rates of 21.65 and 5.0, respectively. Additionally, the overall trend in contact crimes, defined as crimes where victims are the target, such as robberies and assaults, peaked during the beginning of the sample period of this analysis after several years of sustained increases [47]. Contact crimes temporarily decreased between 2016 and 2018, but have suffered large increases thereafter, as shown in Figure 1.1 from a yearly crime report published by SAPS.

1.4 Data

I combine two major sources of data to analyze the impact of changing cash availability on social grant payment dates on crime. These are data on top-level criminal incidents and data on social grant payments. The South African Police Service (SAPS) publishes crime statistics on community reported and police detected serious crimes such as murder, robbery, arson, theft, sexual offenses, and drug related offenses among others. These statistics are at the quarterly level and are aggregated to the province level. SAPS also publishes a yearly report on the change over the year in crime trends. However, my analysis requires data with information on the precise location and date of each crime, in order to determine proximity to social grant paypoints. Through extensive communications with the statistics department of SAPS, I acquired incident-level data for some major categories of crimes that include geographical coordinates for over half of incidents. SAPS provided the data for the province of the Eastern Cape, the second most populous province out of nine provinces. Some crime categories, including murder, arson, stock theft, shoplifting, illegal possession of firearms or ammunition, and alcohol and drug related crimes, were excluded from the data I received. Since this list includes financially motivated crimes that might respond to temporary changes to the liquidity of social grant recipients, I cannot fully examine this mechanism. Similarly, the exclusion of drug and alcohol related crimes means I cannot test for effects related to the consumption of these products. In addition to the geographical information of incidents and their offense type, the data include the date and time. Table 1.2 details the offense types present in the data. A total of 574,163 incidents are reported between April 2014 and March

2021. However, the sample for analysis is restricted to days before the onslaught of the COVID-19 pandemic in March 2020, resulting in 477,441 incidents. Table 1.3 presents tabulations of crimes by category over the sample years. Incidents increase substantially between 2014 and 2015, thereafter remaining at about 80,000 per year.

SAPS collects incident-level data through their central Crime Administrative System (CAS) used by all 1,154 police stations. Police officers open case dockets in the system and record crimes reported by victims, witnesses, third parties, and/or detected through police activity. Officers input all relevant information of the incident into the CAS system. SAPS methodology documents provided through internal communications state that the number of counts associated with each offense are determined after the record of the incident takes place. Given the structure of the data, it is at the offense-level and not at the count-level. As with crime data produced in the US, United Kingdom, and Canada, among others, only the top-level charge is recorded. I proceed in this analysis with the assumption that more violent incidents will be recorded as the top-level charge when an incident includes both violent and non-violent offenses. The data include an "offense ID" variable that is unique for all incidents. In the data, 99.16 percent of observations the 254,129 incidents with coordinate data are unique in date, time, and location. Furthermore, for the 117,442 incidents that do not have coordinate data but include the name of the place where they occurred, 98.59 percent are unique in date, time, and name of place. This strongly suggests the data are at the offense-level, with the top charge, and not at the count level. Due to this structure, we can interpret the counts of violent crimes to accurately reflect the occurrence of violent crime. However, this structure also implies that counts for other crime categories, such as non-violent robberies and thefts, are likely to be under-counted. Therefore, I present all results for only total crimes and violent crimes.

Although only 52 percent of observations contain precise geographical coordinates from SAPS, I am able to recover coordinate data for an additional 46 percent of incidents by implementing geocoding methods using the information on the name of the place, suburb, and district of where the incident took place[2]. This leaves 98 percent of incidents with precise coordinate locations. Incidents without coordinate locations recorded by SAPS do not vary systematically from those with coordinate information. Table 1.4 presents regression results testing for a significant correlation between the likelihood of missing coordinate data and the day of the month and police precinct. As shown in column (1), there is no relationship between day of the month, police precinct, their interaction, and the probability of missing geographical data. For analysis, I aggregate incidents to the police-precinct level. To illustrate the relative size of a precinct, Figure 1.3 visualizes the provinces (administratively similar to states in the US), districts, local municipalities, and finally, the smallest, police precincts. As depicted, precincts constitute geographic areas that are much smaller than the smallest administrative areas, local municipalities, as determined by the federal government. On average, there are 5.96 police precincts within a given municipality, the maximum is 22 in the metropolitan municipality of Buffalo City. Figure 1.4 plots the locations of each police station and depicts a higher concentration of police stations in urban areas.

The issue of crime underreporting plagues all crime data and is likely to be more serious in middle- and low-income countries [48, 49]. The wide literature on crime underreporting states that crimes are more likely to be reported as their severity increases [50–54]. For example, according to the South African Victims of Crime Survey 2016/2017, 56.7 and 51.2 percent of households that

[2]Names of places are mostly the name of the establishment or sometimes the name of the neighborhood where the incident took place.

experienced them reported home robberies and burglaries to the police, respectively. Alternatively, 94 percent of households report motor vehicle theft when experienced. At the individual level, only 34 percent of individuals who are victims of theft of personal property report it to the police. As a consequence, underreporting of crime results in fewer charges and prosecutions than crimes that occurred. Although for several reasons we may not expect underreporting in South Africa to be similar to the US, as an example, in the major metropolitan city of Baltimore, MD, only 44 percent of property crimes and 53 percent of violent crimes are reported [55]. Ultimately, the result of crime underreporting is a systematic under-measurement of crime. While crime underreporting is well documented, there exists no empirical evidence of over-reporting of crime. Therefore, underreporting of crime constitutes non-classical, one-sided measurement error in the dependent variable that leads to estimates representing a underestimate of the true effect [56].

The second major source of data used in this analysis are records of social grants and paypoints. The South African National Treasury published records of all paypoints and their geographical location coordinates in 2015[3]. Data on social grant payment days are widely published across media outlets including the social media pages of SASSA. These data are combined with the incident-level crime data to determine the location of crimes in relation to paypoint locations and police precincts. Importantly, a comparison of paypoint locations between 2015 and 2017, as well as public reports, indicate that paypoint locations did not change over time, that is, until a mass paypoint decommissioning in late 2018. In the Eastern Cape province, there were 2,602 paypoints during the time frame of this analysis. Social grant beneficiaries have the ability to visit any paypoint to retrieve their grant payments. If beneficiaries strategically select what paypoint location they visit, for example, by substituting away from paypoints located in high

[3]Susan Godlonton provided paypoint location data for 2017.

crime areas, then paypoints represent an imperfect proxy for levels of cash disbursement in a given police precinct. However, traveling to alternate paypoint locations would likely result in additional transportation costs that may deter beneficiaries from engaging in this behavior since paypoints are not typically in close proximity to each other. Table 1.5 shows the calculated distances between paypoints throughout the province, by district. This varies from a low of 2.3km in Nelson Mandela Bay to a high of 11.3km in Cacadu. Additionally, most recipients walk to paypoint locations. In a 2012 survey of social grant beneficiaries across two South African provinces, 92 percent of respondents reported living within walking distance to their nearest paypoint location [57].

Lastly, I use auxiliary data on police precinct boundaries to define the main geographical unit of interest, the police precinct, and additional data on climate, population, and municipality demographic characteristics to include as covariates in the estimation models. Police precinct boundaries are published by SAPS and are regularly available on their website. Climate data on daily precipitation and temperature come from the National Oceanic and Atmospheric Administration (NOAA). I use precinct-level population data constructed by Lizette Lancaster and Ellen Kamman of the Institute for Security Studies based on 2011 census population data published by Stats South Africa to determine the number of paypoints per 10,000 residents in a police precinct. I include municipality-level demographic characteristics also from the 2011 census. These are presented in Tables 1.6 and 1.7. These data are not available at the police precinct level, so I match precincts to their respective municipalities, the next largest administrative geographical area. This is an imperfect comparison of precincts along demographic characteristics given that precincts can vary substantially within municipalities, such that a municipality can contain precincts located in both rural and urban areas. The municipality characteristics can be informative on the relative differences between places with differing levels of paypoints. The municipality characteristics include an indicator for whether the municipality is defined as a metropolitan area; mean age; percentage of adults that have completed 12 years of schooling; the unemployment rate; percent white; percent poor; average household size; and others. Paypoint locations are not random; Tables 1.6 and 1.7 summarize and test the difference across municipalities with varying levels of paypoint locations. The results show that places with more paypoints tend to have higher population levels, higher percentages of the population under the age of 15, lower levels of completion of higher education, lower proportions of households in formal dwellings, among other differences.

1.5 Estimation of the Effects of Social Grant Payment Days on Crime for Precincts With/Without Paypoints

To estimate the causal impact of the increase in cash availability that occurs on social grant payment days in the vicinity of paypoints on crime, I employ DiD empirical strategies that exploit the temporal and geographic variation of criminal incidents and social grant payments. First, I estimate a standard two-way fixed effects DiD model where I compare average daily crimes on paydays at the precinct level compared to other days of the month in precincts with and without social grant paypoints. The empirical specification is as follows:

$$\begin{aligned} y_{dpm} = \beta_0 + \beta_1 SG_d + \beta_2 Paypoints_{pm} + \beta_3 \mathbf{SG_d} \times \mathbf{Paypoints_{pm}} \\ + \gamma W_d + M_m + X_d + month_d + weekday_d + day_d + \epsilon_{dpm} \end{aligned} \tag{1.1}$$

Where y_{dpm} denotes average daily crime counts on date d, police precinct p, and municipality

m. SG_d is a binary indicator for monthly social grant payment dates and $Paypoints_pm$ indicates police precincts with paypoint locations. For additional covariates to increase the precision of the estimates, I include W_d, a weather control vector that includes average daily precipitation and maximum temperature at the province (not precinct) level, following the daily crime literature [58,59]. Since weather data is only available at the province level, it is the same for all locations in the data, but it varies by date. M_m includes the municipality demographic characteristics measured in 2011 that do not vary over time, while X_d is a vector of indicators for national holidays and monthly salary paydays, both of which vary monthly and yearly. The fixed effects include $Month_d$, a month by year fixed effect for the 53 months between April 2014 and September 2910 and $weekday_d$, a day-of-week effect for the seven days of the week. Day_d is a quadratic day-of-month control that ranges from one to thirty-one that is not collinear with the daily weather controls since it does not capture unique dates, but only day of the month. Police precinct population data is not available for two precincts, resulting in a sample with 195 out of 197 precincts. The parameter of interest in this model, β_3, captures the effect of the presence of paypoint locations disbursing cash on social grant payment days, therefore of increasing local cash availability and the expected returns to criminal activity. Precincts with zero paypoints constitute 11 percent of total precincts. Figure 1.5 shows the variation in the distribution of the number of paypoints by precincts. Most precincts have less than 20 paypoints, with the majority housing less than 10 paypoints. There are also precincts, in highly urban settings, with over 100 paypoints.

To identify a causal estimate of the effect of social grant paydays on daily average crime using this DiD specification, it must be that daily crimes on non-social grant paydays do not experience differential trends in precincts with paypoints compared to those without. Additionally, there must not be any potential time-varying confounders that affect precincts with and without paypoints differently. While it is impossible test for these differential trends since we do not observe the counterfactual, namely, the crimes that would have occurred in treated precincts if they had not had paypoint locations, I test for a modified parallel pre-trends assumption to the standard DiD parallel pre-trends. Because social grant payment days occur monthly, these repeated treatment periods make it impossible to conduct the usual analysis of trends in the outcome variable before an event, such as a policy change. Instead, I confirm that days surrounding the social grant payday exhibit similar patterns of daily average crime in precincts with and without paypoint locations. Figures 1.7 and 1.8 show residualized trends in average daily crime, and Table 1.8 tests for differential modified pre-trends. Note that Day 0 on the x-axis of Figure 1.7 corresponds to the social grant payment date of each month and this figure shows average crime patterns across the entire month. Figure 1.8 shows the average crime trends for only the middle two weeks of the month to abstract from the potential effects of the grants in the surrounding days - at the end of the month leading up to payment and then throughout the first week following payments. The figures show clear spikes in all crimes and violent crimes, on and around social grant paydays. In particular, we observe increases in crimes about five days prior to the social grant payments. This corresponds to about the 25th day of the month, the unofficial but widely adopted salaried worker's payday. All empirical specifications include a control for the salary payday to account for its potential effects on crime, The trends on non-paydays are the same in places with and without paypoints. Tables 1.8 and 1.9 show the results of a test using the main specification in Equation 1.1 but excluding social grant paydays from the data, and instead reporting the coefficient on the interaction term [$\textit{Day of Month}_d \times \textit{With Paypoints}_{pm}$]. I also include a test solely for the week following social grant payment week in Table 1.10. Across offense category, not only are the coefficients close to zero, they are also statistically insignificant with large p-values. These results suggest we cannot reject the null of the modified parallel pre-trends assumption.

1.5.1 Results

Table 1.11 presents the results of estimating Equation 1.1 considering the social grant payment day as the date of treatment, as well as alternative specifications testing effects of the day before and the day after payment day. The parameter of interest, β_3 on the interaction term, captures the effects on daily average crime for comparing police precincts with no paypoints to those with an average of 5.25 paypoints per 10,000 residents. The results are presented in Table 1.11 where columns (1) and (3) present specifications including municipality-level demographic controls and columns (2) and (4) remove these covariates and instead include police precinct fixed effects to capture unobservable time-invariant differences across precincts. The main effect of having paypoints in a precinct can no longer be estimated for these specifications as paypoint locations are static across time. Looking at the second panel of Table 1.11, the estimates suggest a positive but close to zero effect of increased cash availability on social grant payment days on average daily crimes across specifications. Point estimates for all crimes and violent crimes are small relative to mean daily counts. For example, $\beta_3 = 0.035$ represents only a 1.45 percent increase from the control group mean.

Focusing on the first panel of Table 1.11, though imprecisely estimated, we observe a decrease in crime on the day prior to social grant payments. For example, all crime decreases on average by 6.77 percent, with a p-value of 0.077. Estimates in the third panel for the day after grant payments show an increase in all crimes and violent crimes. Moreover, the estimated increase in all crime is substantial, with a coefficient of $\beta_3 = 0.220$ representing an 9.14 percent increase in average daily crimes from the control group mean of 2.40. Testing the equality of the estimated coefficients for all crime on the interaction terms for the prior and next days compared to the payment day results in p-values close to 0.10 as shown in the fourth panel of the table. For violent crimes, we cannot reject the equality of the estimated coefficients. However, the estimated effects in columns (1) and (2) for the prior day are statistically significantly different from those on the next day to social grant payments for all crime, with p-values of 0.00 for both. These results may be the result of a displacement in crime. I test whether the sum of the estimated coefficients for prior, pay, and next day of social grant payments is equal to zero, and for all specifications and crime categories, the resulting p-values are upwards of 0.28. Thus, the observed decrease in crime prior to the grant payments is offset by the increase in crime following the payments.

1.6 Estimation of the Effects of Social Grant Payment Days on Crime for Police Precincts with Above/Below Median Number of Paypoint Locations

The second main empirical strategy re-estimates Equation 1.1 but defines treatment and control precincts by the number of paypoints located within each precinct to compare crime in places with above and below than the median number of paypoint locations of 2.48 per 10,000 residents:

$$\begin{aligned} y_{dpm} = \alpha_0 + \alpha_1 SG_d + \alpha_2 High\_PP_{pm} + \psi \mathbf{SG_d} \times \mathbf{High\_PP_{pm}} \\ + \gamma W_d + M_m + X_d + month_d + weekday_d + day_d + \epsilon_{dpm} \end{aligned} \tag{1.2}$$

All variables are defined as in Equation 1.1, except now SG_d is interacted with $High\_PPs_pm$, an indicator for whether a precinct has above 2.48 paypoints, the median number of effective paypoints per 10,000 residents. The parameter of interest, ψ, captures the effect of moving from an average of about one paypoint in places with low numbers of paypoints to almost nine in places with high numbers of paypoint locations.

As before, the identification of the parameter of interest ψ relies on treatment and control precincts experiencing similar changes in average daily crimes across offense types on days without social grant payments. The residualized trends showing the modified pre-trends assumption in average crimes are presented graphically in Figures 1.9 and 1.10 for the entire month and the two weeks following social grant payment week. The results of the tests of the depicted patterns in average daily crimes are presented in Tables 1.12 and 1.13 using the same covariates and fixed effects as the main specifications. Table 1.14 tests these patterns solely for the week following social grant payment week. Again, we observe large spikes on and around social grant payment dates for all crimes and violent crimes. Moreover, we still fail to reject the null hypothesis for parallel trends on non-payment days.

I also estimate the potential effects of an additional paypoint location in a given police precinct on crime through a Generalized DiD specification that exploits the variation in the number of paypoint locations in a precinct and the timing of the social grant payments:

$$\begin{aligned} y_{dpm} = \eta_0 + \eta_1 SG_d + \eta_2 No.Paypoints_{pm} + \delta \mathbf{SG_d} \times \mathbf{No.Paypoints_{pm}} \\ + \gamma W_d + M_m + X_d + month_d + weekday_d + day_d + \epsilon_{dpm} \end{aligned} \quad (1.3)$$

Where $No.Paypoints_{pm}$ denotes the number of paypoints per 10,000 residents, excluding precincts without paypoints. The parameter of interest δ captures the effect of an additional paypoint location on average daily crimes on social grant payment days compared to other days. All other variables are as previously defined. The identification of δ, as with the previous model, relies on the assumption that changes in average daily crimes in places with lower numbers of paypoint locations serve as an adequate counterfactual for changes in average daily crimes for places with higher numbers of paypoint locations. I explored this assumption above, comparing precincts with high and low numbers of paypoints, in Figures 1.9 and 1.10, and tested the trends on non-payment days in Tables 1.12, 1.13, and 1.14. In addition, Table 1.15 presents the results from testing for differential trends in average daily crimes using the main specification in Equation 1.3 but excluding social grant paydays from the data, as before, for the middle two weeks of the month. From the results, I am unable to reject the null of the modified parallel pre-trends on non-social grant payment days for this specification as well.

1.6.1 Results

Table 1.16 presents results from the estimation of Equation 1.2, where the relevant comparison is now the change from an average of 0.84 paypoints per 10,000 residents in precincts with a low number of locations, to 8.52 paypoints in precincts with a high number of locations. Consequently, compared to the previous results that estimated the effect of having at least one paypoint, this model estimates the parameter of interest at a different margin. As in the previous tables, the odd numbered columns in Table 1.16 present specifications with municipality demographic controls while the even numbered columns present specifications without these controls but with police precinct fixed effects. The results in the second panel of the table differ from the prior analysis in Table 1.11 in that we no longer observe an increase in crime the day following payments, but we do observe an increase on the social grant payment date. The estimated $\psi = 0.191$ shows the effect of moving from 0.84 to 8.52 average paypoints to be a 6.24 percent increase from the control group mean.

For the days surrounding payment, results in the first panel of Table 1.16 show a statistically significant decrease in all crime the day prior to social grant payments when comparing places with above and below median number of paypoints. Additionally, the point estimate of $\psi = 0.151$ for all crime is a 4.93 percent decrease from the mean in precincts with below median paypoints. The magnitude of this effect is slightly smaller than that from Table 1.11, but more precisely estimated. The change in crime for both the prior and pay day of social grants is concentrated in violent crimes. Additionally, the coefficients on the interaction terms across the payment days are significantly different, highlighting the potential for different mechanisms at play on the various days. These results are consistent with perpetrators potentially delaying their labor supply of crime until the widely publicized cash grant paydays, an anticipation effect, and increasing their labor supply of crime on payment days consistent with the loot effect. Media reports detail how it is a common practice for grant beneficiaries to spend a substantial portion of their cash grants at the shops in which some of the paypoints are located [2]. Thus, if the mechanism is the loot effect of perpetrators stealing cash and purchased goods from transfer beneficiaries, then we expect an immediate effect in the vicinity of the paypoints since concentration of loot will be highest at and around the time of disbursement. However, as with the previous results in Table 1.11, the observed effects may be due to a displacement in crime. Repeating the test of whether the sum of the estimated coefficients on the interaction terms for prior, pay, and next day of social grant payments results in p-values upwards of 0.72 for all specifications and categories of crime.

Finally, Table 1.17 shows the estimation results from the generalized DiD specification from Equation 1.3. All coefficients, across all specifications, days around payment, and category of crime are close to zero effects. However, as with the results in Tables 1.11 and 1.10, we still observe a possible decrease in crime the day before social grant payments. The point estimates are virtually identical across the inclusion of municipality demographic characteristics or police precinct fixed effects. Testing whether the sum of the estimated coefficients on the interaction terms is equal to zero, the resulting p-values for columns (1) and (3) are 0.08 and 0.01, respectively, suggesting that although the coefficients are close to zero, the effects across the days are not offset by each other. Ultimately, these results suggest that small increases in the number of paypoint locations in a given police precinct have almost no effect on crime around social grant payment dates.

1.7 Robustness Checks

The results presented above are robust to a number of alternative treatment definitions and sample restrictions. First, I extend the main analyses by estimating the effects of social grant payments not only on the surrounding days of the payment date, but to the week following payments. This allows me to investigate the potential persistence of the main results and allows for other channels potentially affecting daily crime. It could be that estimating the effects of temporary changes in cash availability on crime is better examined at the weekly crime count level, as opposed to at the daily level, to investigate the potential persistence of the main results, or the possibility that the changes in the incentives or opportunities to commit crime materialize along a longer time frame than a single day. Because South Africa remains a cash driven society, we can expect that recipients are not typically depositing their grant money, especially given that they had the option of direct deposit over cash disbursement and yet most recipients opted for cash. Table 1.18 presents results from estimating the effects of social grant cash payments at the weekly crime level for places with paypoints compared to those without:

$$y_{wpm} = \beta_0 + \beta_1 Week_w + \beta_2 Paypoints_{pm} + \beta_3 \mathbf{SG\_Week_w} \times \mathbf{Paypoints_{pm}} \\ + \gamma W_w + M_m + month_w + \epsilon_{wpm} \qquad (1.4)$$

Where now the outcome variable y_{wpm} is measured as a weekly sum, SG_Week_w is an indicator for the week beginning with the social grant payday, and W_w and M_m are weekly averages of the weather and municipal covariates. Allowing for analysis at the weekly level expands the duration for which social grant cash payments can affect crimes beyond a single day, potentially capturing the spillover effects on the subsequent days following the cash disbursements. The results are presented in Table 1.18 and quantitatively small and statistically insignificant effects on weekly crime counts on social grant payment weeks in precincts with paypoints compared to those with an average of about five.

I repeat the exercise of aggregating crimes to the weekly level for the model comparing precincts with high and low numbers of paypoint locations:

$$y_{wpm} = \alpha_0 + \alpha_1 Week_w + \alpha_2 High\_PP_{pm} + \psi \mathbf{SG\_Week_w} \times \mathbf{High\_PP_{pm}}$$

$$+ \gamma W_w + M_m + month_w + \epsilon_{wpm} \quad (1.5)$$

Results are presented in Table 1.19. The point estimates on ψ for all crimes and violent crimes are all statistically significant and represent substantial increases from the control group mean. For example, at the weekly level, ψ = 1.093 represents an increase of 12 percent in average total weekly crimes on social grant payment weeks in high paypoint precincts compared to low paypoint precincts. These results are consistent with the previous results where we observed increases in crime on the day of or following social grant payments.

Secondly, to further investigate the driving forces of these results, we may believe that more urban regions, such as those with the highest levels of development and population, vary in important ways that could be driving the observed results. I re-estimate the main models from Equations 1.1 and 1.2 dropping the largest two cities in terms of population and urbanization in the Eastern Cape, namely Port Elizabeth and East London. It is possible that in these areas with, for example, more drinking establishments, or other places and activities that are complementary to criminal behavior, perpetrators behave differently than those in areas with a lower stock of these complementary goods and services. A recent analysis in 2019 mapping drinking establishments across urban and rural settings in South Africa confirm much higher concentrations in urban settings [60]. Port Elizabeth and East London contain 25 percent of daily observations in our main sample. Interestingly, although these places have the highest population levels, they do not house the highest numbers of paypoint locations. The average number of paypoints in precincts in Port Elizabeth and East London is 5.25, compared to 5.82 for those outside of these cities. Similarly for precincts with higher relative number of paypoints, in the restricted sample the average number of paypoint locations is also slightly higher at 8.59 compared to 8.52, awe well as for those with a relatively low number at 0.99 compared to 0.84. However, importantly, the difference between high and low paypoint precincts remains almost identical between the samples, at 7.6 effective paypoints in the restricted sample, and 7.68 in the full sample.

Results are presented in Tables 1.20 and 1.21. Beginning with the comparison of precincts with and without paypoint locations, the estimated coefficients confirm the previous findings of decreases in crime the days before social grant payments and increases on the days of payment and the following days. While the estimates in the first and second panels are similar in magnitude to the main results in Tables 1.11 and 1.16, they are imprecisely estimated. For example, the second panel shows an increase in crime on social grant payment date of 0.151, yielding a p-value of 0.073. The point estimates in the third panel for the day following payments are also similar in magnitude to the previous results, for example, $\beta_3 = 0.256$ compared to $\beta_3 = 0.220$ in Table 1.11. With the magnitudes of the coefficients mostly unchanged, it is clear that the loss of statistical significance comes only from the increased standard errors from the drop in sample size. Next, for the comparison of precincts with above and below median number of paypoints, Table 1.21 also shows similar patterns as the previous results. We observe a decrease in crime the day prior to payments, and an increase in crime on the day of payments, both with similar magnitudes as Table 1.16. Additionally, the coefficients on the interaction terms on and surrounding social grant payment day are significantly different from each other, as the near-zero p-values in the fourth panel show.

Another potential driver of the main results using the variation in the number of paypoint locations could be those places with the highest number of paypoints, or conversely, those with the lowest number of paypoints. There are many possible reasons why precincts at the extremes of the number of paypoints distribution may be somehow different than the rest. For example, precincts without paypoints are located in more remote areas of the province. It may be more difficult to reach the mobile and temporary paypoints locations in these areas, and for similar reasons, local crime patterns may differ. I present the results of re-estimating Equation 1.2 excluding precincts with zero paypoints in Table 1.22 and excluding precincts with the top 5 percent number of paypoints in Table 1.24. The results in both tables are similar to those in Table 1.16. The magnitude of the effects on payment days are almost identical, with the main results estimating an effect of $\psi = 0.191$, compared to $\psi = 0.186$ and $\psi = 0.193$ for the zero and top paypoint exclusions, respectively. The results also confirm that the effects across the social grant payment days are substantially different from each other, resulting in p-values smaller than 0.001 across all specifications and crime categories. I also re-estimate the generalized DiD specification in Equation 1.3 and present the results in Table 1.23. The results remain quantitatively unchanged, with estimated effects close to zero but with similar patterns of decreases on the days before social grant payments and increases on the day of payments.

Lastly, another feature of the South African context is that many workers, including government employees, are paid their salaries, via direct deposits, cheques, or cash, on a monthly basis on the 25th day of the month [61–63]. Additionally, as with the social grant payments, retail stores experience spikes in goods purchased immediately following the salary payments, resulting in a similar loot effect as a potential mechanism [63]. Also similar to the cash grant payments, they vary each month, for example, if the 25th falls on a weekend or national holiday then the following business day becomes the payday. I re-estimate the main models in Equations 1.1 and 1.2 using salary payment days instead of social grant payment days and present the results in Tables 1.25 and 1.26. When comparing precincts with and without paypoints, we observe increases in crimes both on the salary payment day as well as the days surrounding the payment. I find different results in Table 1.26 for the comparison of precincts with high and low numbers of paypoints. We now observe no effect on total crime on salary payment days and the coefficients across the days surrounding the payments are not substantially different from each other. The corresponding p-values of a test of the estimates on the interaction terms yields values upwards of 0.48. The result from this falsification test show effects of cash in police precincts with paypoints even when there is no evidence that cash was distributed disproportionately from salary payments in those places. These results are unexpected and the subject of future analysis.

1.8 Discussion

In this analysis, I document an anticipation effect of decreased crime on the days prior to temporary increases in cash availability through social grant payments around their disbursement locations. I also estimate small potential increases in crime on and the following day of social grant payments. To measure these effects, I study the large cash transfer program in South Africa that prior to 2018 disbursed benefits mostly in cash on regularly scheduled monthly intervals at physical payment sites throughout the country. I use incident-level crime data provided by their national police service that includes precise geographical information for about half of all incidents and used qualitative information about the remainder of the incidents to determine their precise locations. Combining these data, I employ two two-way fixed effects DiD specifications exploiting the temporal and geographic variation in cash disbursements as well as incidence of crime, and a generalized DiD specification designed to exploit the variation in the distribution of paypoint locations across police precincts.

Interpretation of the results in this analysis suggest that while much evidence exists in other settings showing large effects of changes to cash availability in communities on crime, it is in fact surprising that we do not observe larger or more consistent effects in the context of South Africa's social grant program. A limitation of this analysis lies in the structure of the crime data as top-level charges that potentially masks the true nature of non-violent crime reporting and therefore prevents analysis of some of the most likely affected categories of crime given the context. Additionally, the number of paypoints as a measure for the rate of temporary change in cash availability on social grant paydays is an imperfect proxy and data on number of beneficiaries by police precinct along with the rates at which beneficiaries collect their benefits in cash on the first day they become available would increase the precision of the estimates and alleviate potential measurement error in the independent variables.

An avenue for potential future research is the examination of the mass paypoint decommissioning in the second half of 2018 when officials closed almost 80 percent of paypoint locations, made the South African post offices available for social grant payments, and organized the transition to EBT payments. This policy change likely decreased cash availability on social grant paydays and would allow for the estimation of the effect of cash on crime along the opposite direction. This possibility, and likely several other avenues of analysis, could help shed light on the effects of cash on crime.

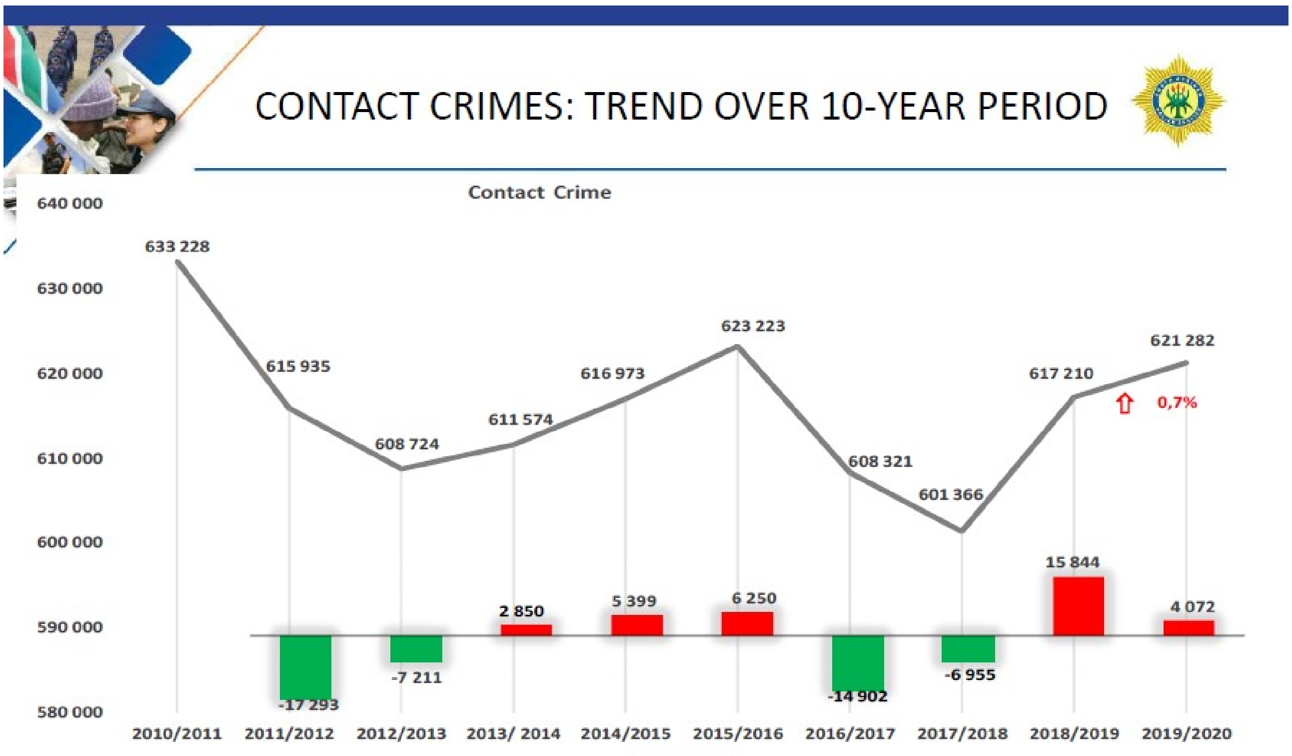


Figure 1.1: Trends in Contact Crimes in South Africa

Source: South African Police Service (SAPS) Crime Statistics Presentation: Crime Situation in Republic of South Africa Twelve (12) Months April 2018 to March 2020

Notes: Figure shows trends in contact crimes, defined as murder, attempted murder, assaults, robberies, and sexual offenses between 2010 and 2020.

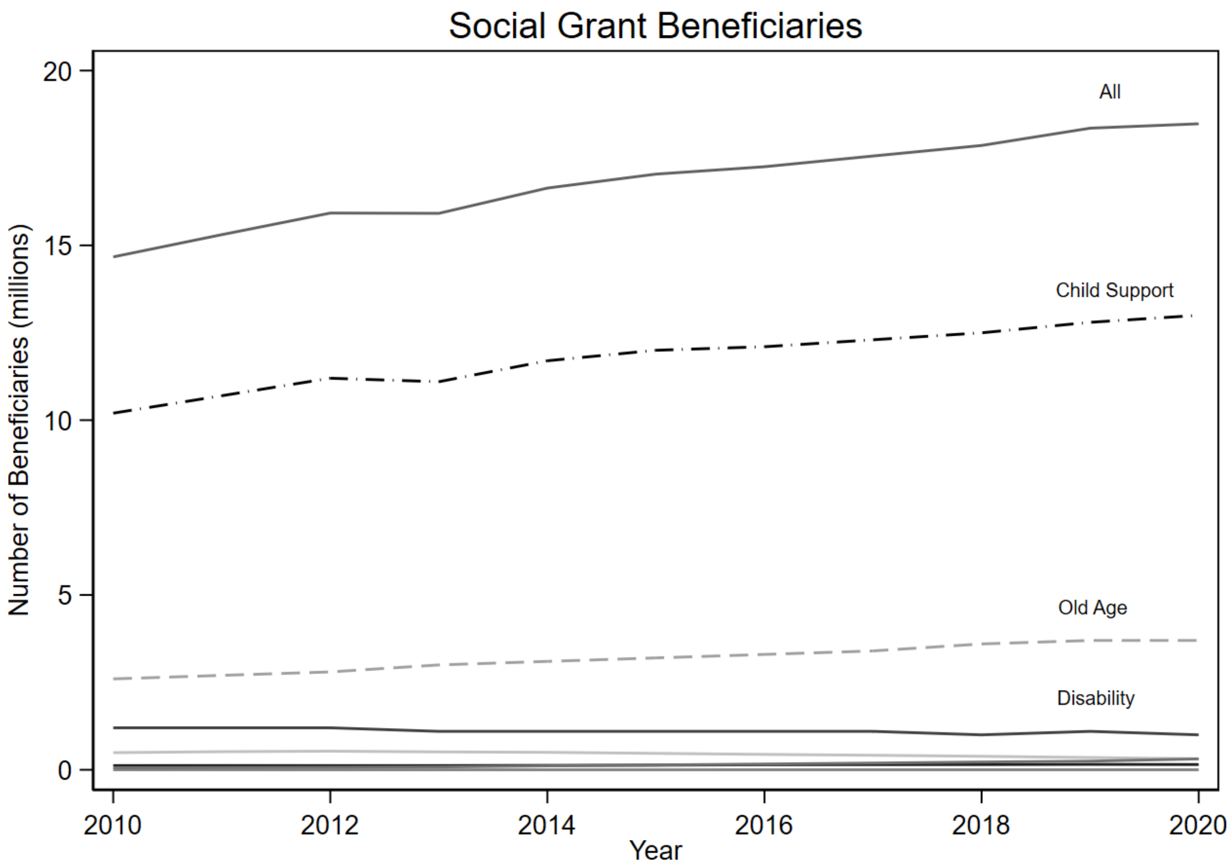
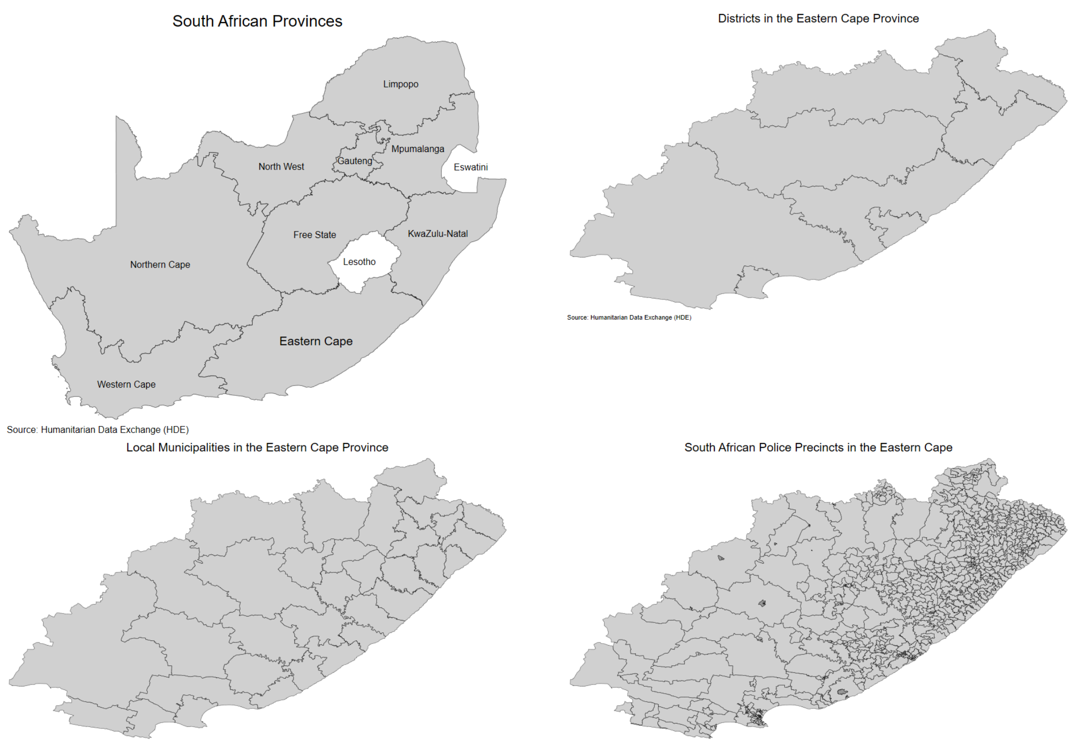


Figure 1.2: Social Grant Beneficiaries in South Africa

Notes: Data from South Africa Social Security Agency (SASSA) yearly social grant reports for 2010 to 2020. Figure shows the number of beneficiaries, in millions, for all six grants: child support grant, old age grant, disability grant, war veterans grant, foster care grant, and care dependency grant.

Figure 1.3: South African Geographic Boundaries



Notes: Shapefile data from Statistics South Africa (Stats SA) and police boundary shapefiles from the South African Police Service (SAPS). Figure shows the various administrative boundaries in South Africa.

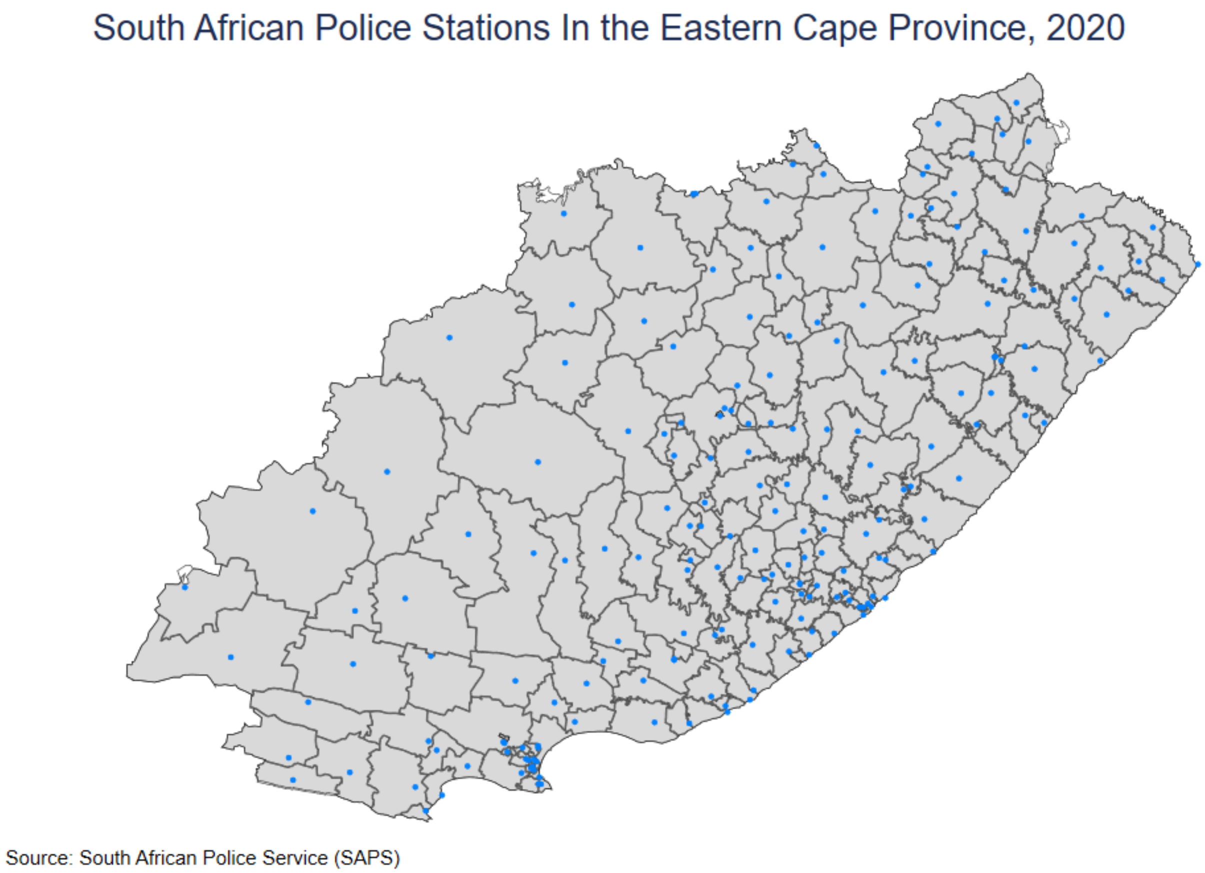


Figure 1.4: Police Stations in South Africa, 2020

Notes: Data on provincial boundaries are from Statistics South Africa (Stats SA) and police station locations are from the South African Police Service (SAPS). Figure maps the locations of police stations throughout South Africa.

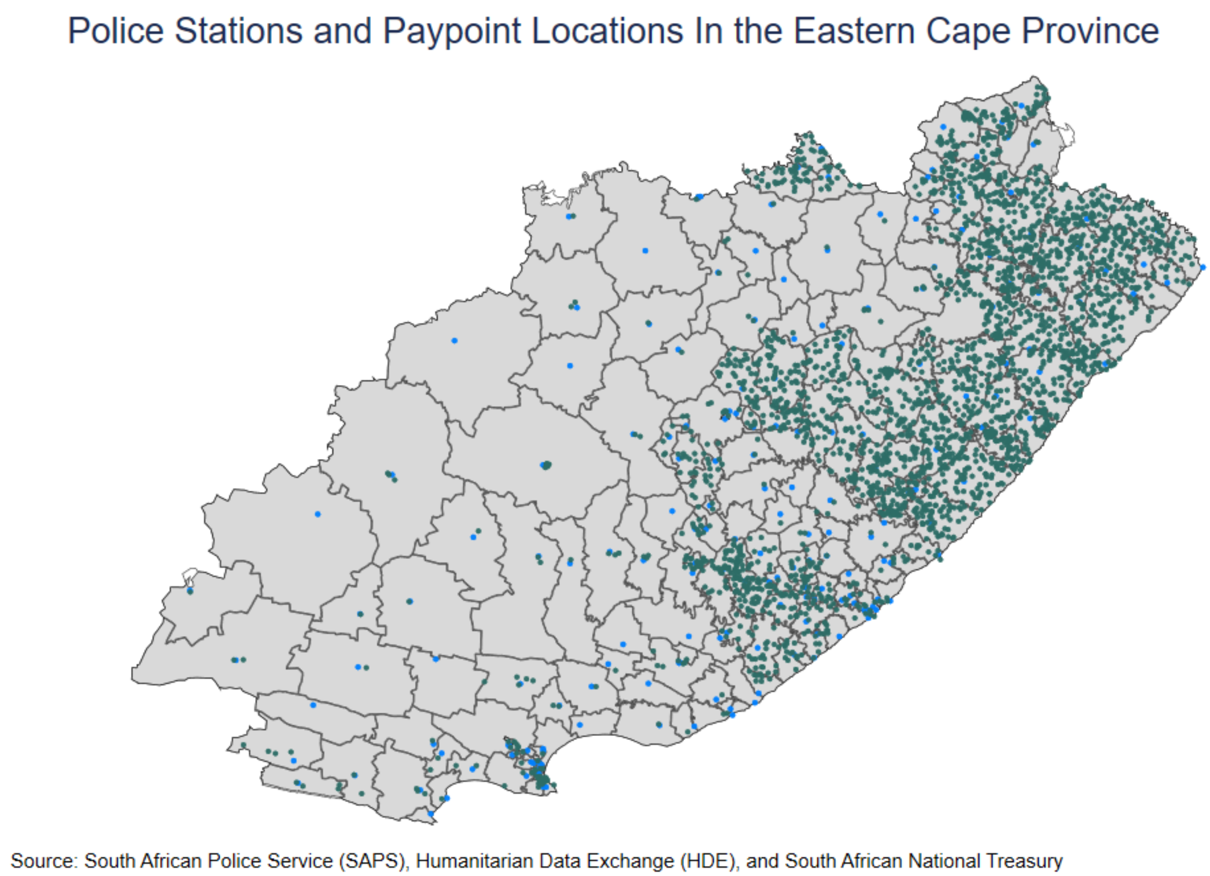


Figure 1.5: Paypoints in South Africa, 2017

Notes: Figure maps the locations of paypoint stations in green, police stations in blue, throughout police precincts in the Eastern Cape province of South Africa.

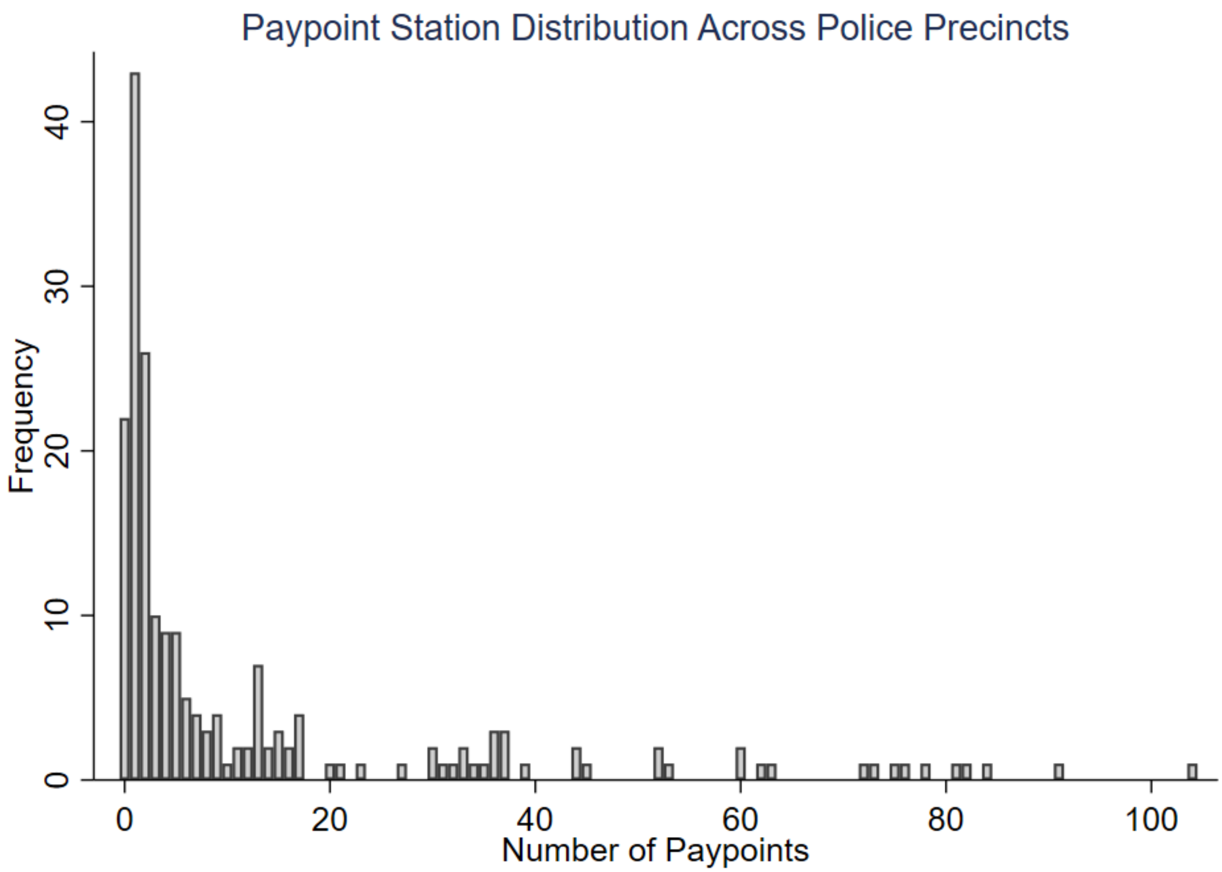
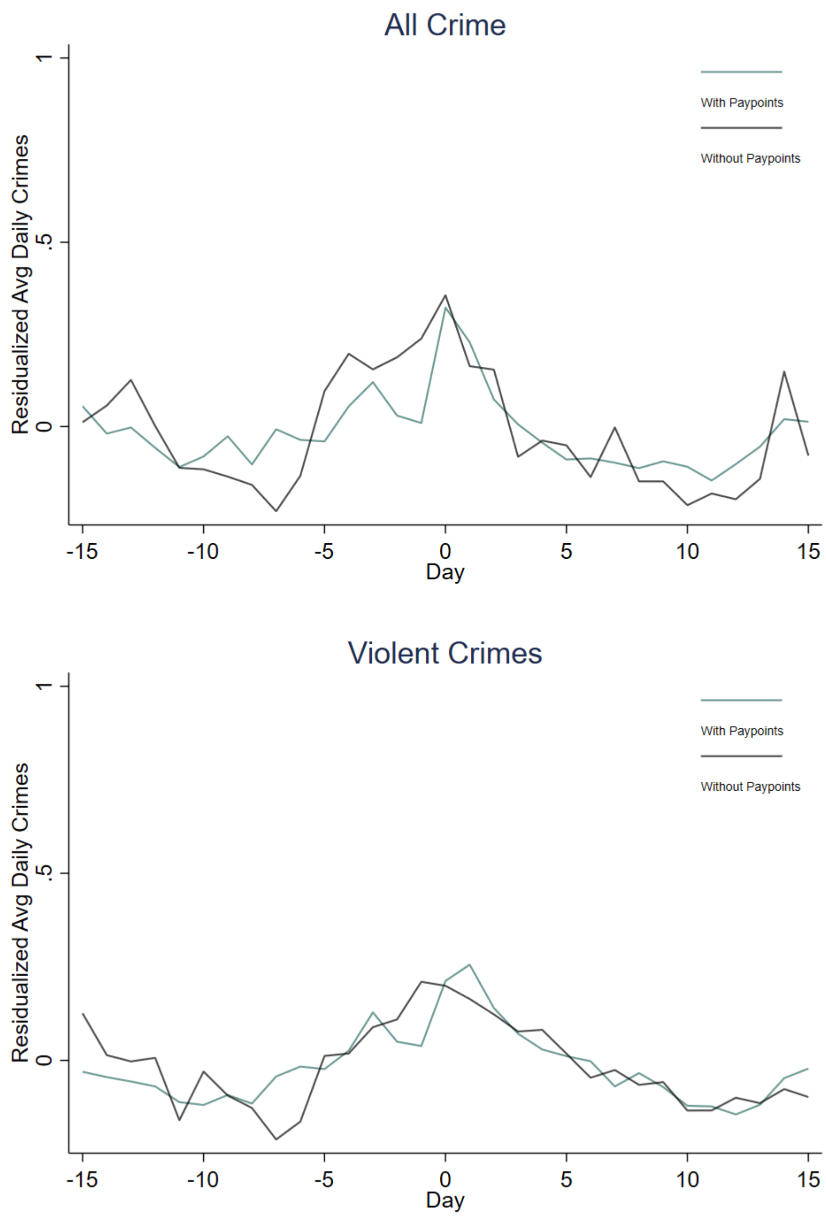


Figure 1.6: Distribution of Paypoints

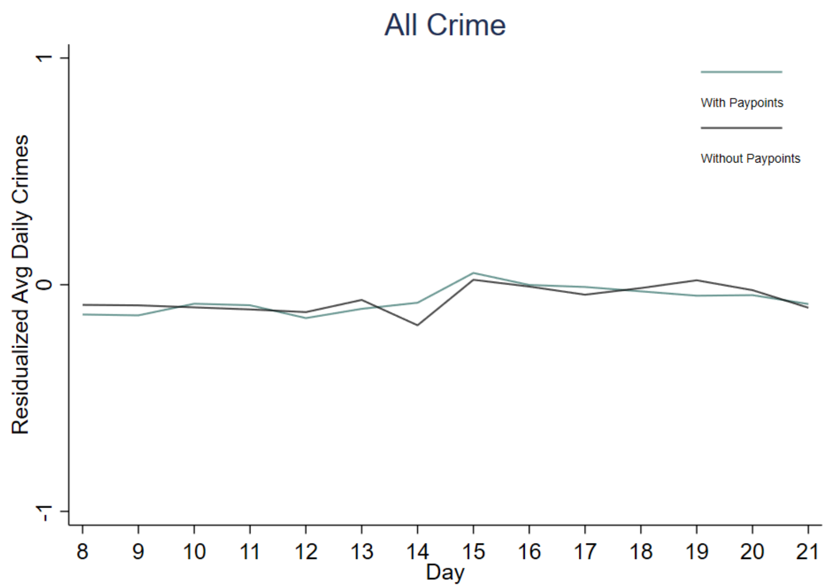
Notes: Data on police precinct boundaries are from the South African Police Service (SAPS) and paypoint station locations are from the National Treasury. Figure shows the distribution of paypoint locations across precincts.

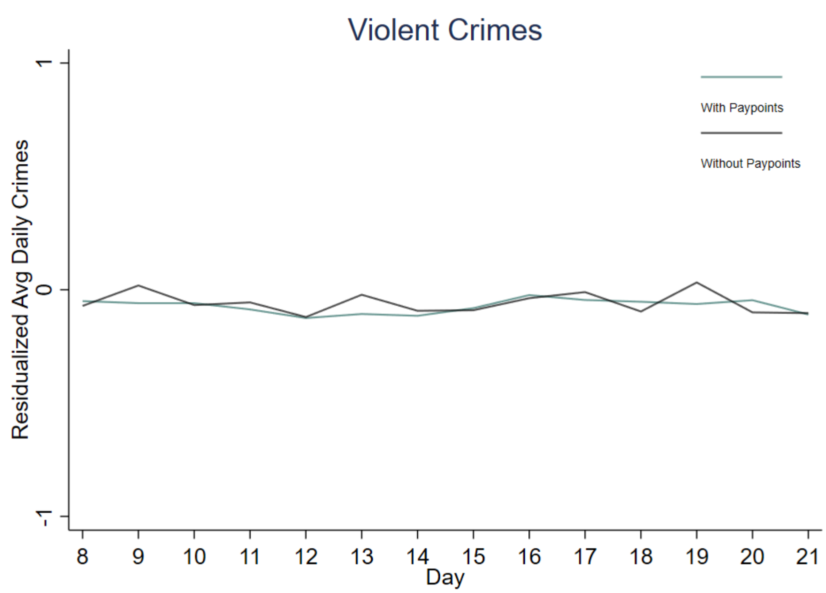
Figure 1.7: DiD With/Without Paypoints: Residualized Avg Daily Crime Trends



Note: Day 0 denotes social grant payment dates. Incident-level data are from the South African Police Service (SAPS). Figures show residualized trends in average daily crimes for all crimes and violent crimes for each day of the month for police precincts with and without paypoint stations.

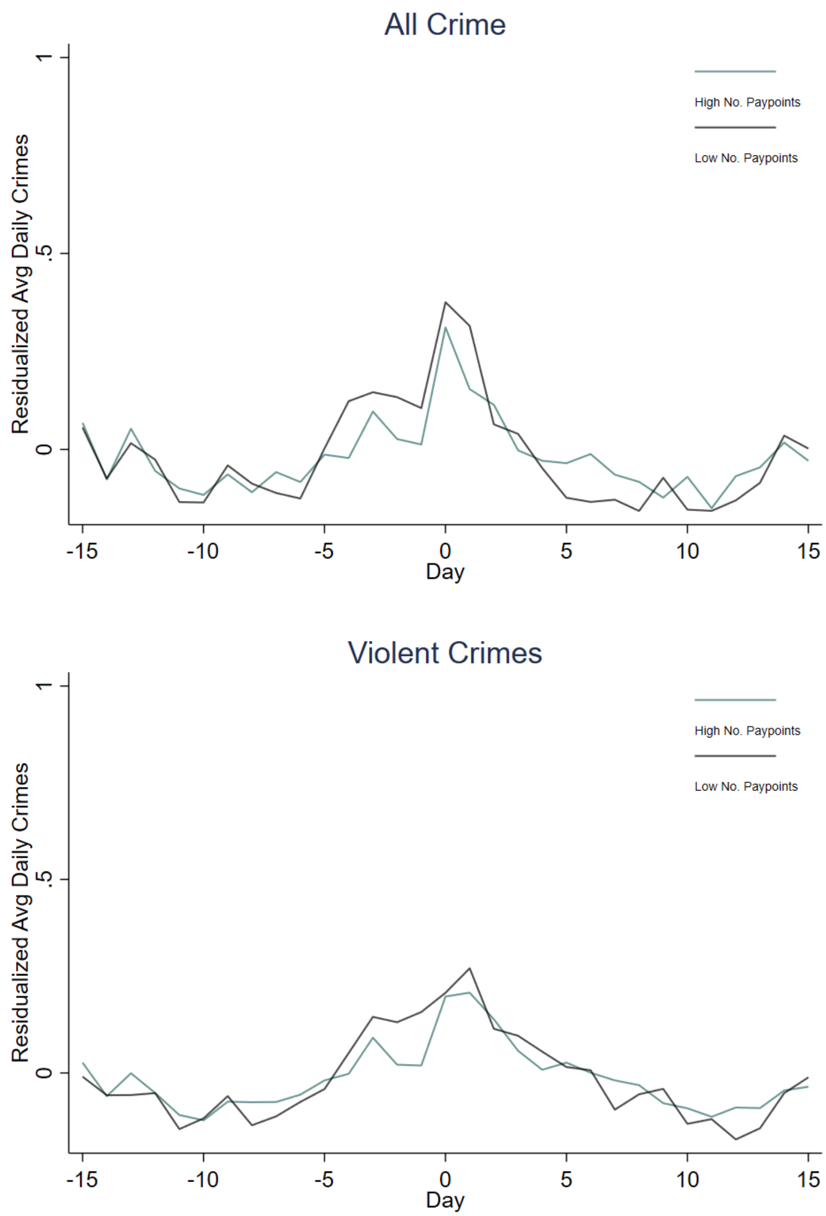
Figure 1.8: DiD With/Without Paypoints: Residualized Avg Daily Crime Trends for Middle of the Month





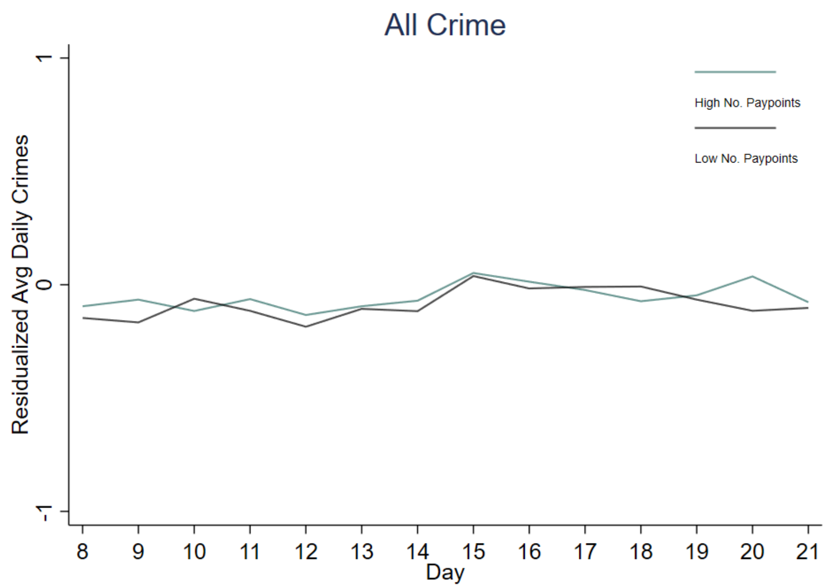
Note: Day 0 denotes social grant payment dates. Incident-level data are from the South African Police Service (SAPS). Figures show residualized trends in average daily crimes for all crimes and violent crimes for each day of the month for the middle two weeks of the month across police precincts with and without paypoint stations.

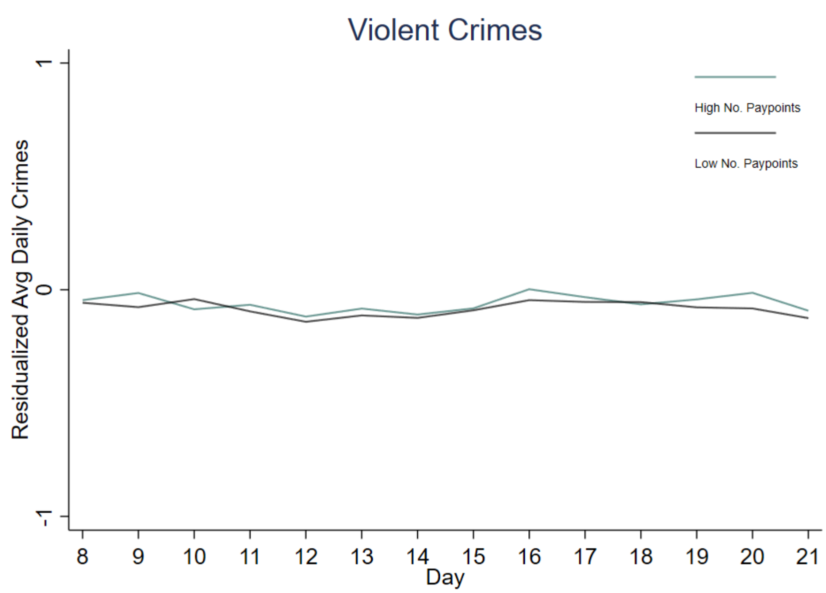
Figure 1.9: DiD: High/Low Paypoints: Residualized Avg Daily Crime Trends



Note: Day 0 denotes social grant payment dates. Incident-level data are from the South African Police Service (SAPS). Figures show residualized trends in average daily crimes for all crimes and violent crimes for each day of the month for police precincts with above and below median number of effective paypoints per 10,000 residents.

Figure 1.10: DiD High/Low Paypoints: Residualized Avg Daily Crime Trends for Middle of the Month





Note: Day 0 denotes social grant payment dates. Incident-level data are from the South African Police Service (SAPS). Figures show residualized trends in average daily crimes for all crimes and violent crimes for each day of the month for the middle two weeks of the month across police precincts with above and below median number of effective paypoint stations per 10,000 residents.

Table 1.1: Value and Prevalence of Social Grants

	Old Age Grant			Child Support Grant			Disability Grant		
Year	Value	Beneficiaries	Expenditure (Billions)	Value	Beneficiaries	Expenditure (Billions)	Value	Beneficiaries	Expenditure (Billions)
2014	1,350	3.1m	49.1	315	11.7m	43.7	1,350	1.1m	18.8
2015	1,415	3.2m	53.1	330	12.0m	47.3	1,415	1.1m	19.2
2016	1,505	3.3m	58.3	355	12.1m	51.6	1,505	1.1m	19.9
2017	1,600	3.4m	64.1	380	12.3m	55.9	1,600	1.1m	20.9
2018	1,695	3.6m	71.0	405	12.5m	60.6	1,695	1.0m	22.0
2019	1,780	3.7m	77.0	425	12.8m	65.0	1,780	1.1m	23.1
2020	1,860	3.7m	83.1	445	13.0m	84.9	1,860	1.0m	24.4

Notes: Figures are nominal values in South African currency, Rands. Data are from the South African National Treasury yearly budget reports. In 2020, recipients of the old age and disability grants must have earnings less than R86,280 if single and R1,227,600 if married, while child support grant recipients must earn less than R53,400 and R106,800, respectively. Recipients over 75 years of age receive an additional R20 in their old age grants.

Table 1.2: Offenses Provided by SAPS

Incident Type	
Financial	***Sexual***
Burglary at Residence	Rape
Burglary at Business	Compelled Rape
Robbery at Residence	Sexual offenses with children aged 16 or less
Robbery at Business	Sexual offenses with mentally disabled persons
Robbery with Firearm	Homosexual offenses
Robbery of Bank	Prostitution or keeping a brothel
Robbery of Cash in Transit with Firearm	Compel Persons To Witness Sexual Act
Carjacking	Public Indecency/Indecent Behaviour/Exposing
Hijacking - Truck	Trafficking In Persons For Sexual Purposes
	Bestiality
Assault	Incest
Common Assault	Use of children or mentally disabled persons
Assault (Grievous Bodily Harm)	in pornography
Sexual Assault	
Compelled Sexual Assault	***Other***
Compelled Self- Sexual Assault	Abduction of minor
Indecent Assault	

Notes: The categories above are as provided by the South African Police Department (SAPS). Compelled crimes are those where one individual forces another individual to commit the crime. The aggregated categories of financial and assaults follow the FBI's Uniform Crime Report (UCR) classifications with the exception of robberies included as a financial crime instead of a violent crime. Sexual offenses include sexual violation or exploitation, grooming, ”flashing”, and forced witness. SAPS defines any non-heterosexual acts as criminal.

Table 1.3: SAPS Yearly Crime Counts, Eastern Cape Province

Year	Robberies & Burglaries	Violent	Sexual	Motor	All Crime
2014	27,532	26,317	6,505	629	60,606
2015	35,014	33,570	8,300	923	77,181
2016	37,982	34,758	7,721	1,025	80,803
2017	37,896	34,165	8,029	1,121	80,482
2018	38,072	36,341	8,158	1,168	82,884
2019	36,984	36,709	8,365	1,200	82,406
2020	5,973	5,807	1,259	183	13,079
Total	219,453	207,667	48,337	6,249	477,441

Notes: South African Police Service (SAPS) provided incident-level crime data. Crime categories are not mutually exclusive as violent crimes include sexual assaults along with all other assaults. Motor crimes include carjackings, hijackings, and robberies of cash-in-transit commercial vehicles. Sexual crimes include all crimes of a sexual nature including rape, indecent exposure, incest, among others. All crimes include the listed categories as well as abductions. The year 2020 shows counts until April 1, 2020.

Table 1.4: Correlates of Missing Geographic Coordinates

	Missing Geographical Coordinates (1)
Day x Police Precinct	0.0000
	(0.0000)
Day	-0.0002
	(0.0003)
Police Precinct	-0.0001
	(0.0001)
Adj. R-squared	0.003
N	477,438

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

Notes: Observations at the incident level. Sample includes incidents occurring between April 2014 and March 2021. Standard errors are presented in parentheses and are clustered at the police precinct level. Estimates show results from a regression of an indicator for missing geographical coordinate data on the interaction between day of the month, ranging from 1 to 31, and police precinct unique identifiers, and the main effects.

Table 1.5: Distance (km) Between Pay Point Stations, Eastern Cape 2017

District	Nearest	2nd Nearest	3rd Nearest	4th Nearest	5th Nearest
Alfred Nzo	2.8	4.1	5.9	6.8	7.9
Amathole	2.9	4.4	5.5	6.4	7.4
Buffalo City	4.5	5.8	6.8	7.9	8.8
Cacadu District	11.3	20.1	32.3	41.2	45.7
Chris Hani	3.2	6.0	7.3	8.4	9.8
Joe Gqabi	5.0	8.4	10.4	12.4	14.8
Nelson Mandela Bay	2.3	2.9	3.3	3.4	4.2
O.R. Tambo	2.7	4.2	5.3	6.7	7.6
All	4.3	7.0	9.6	11.7	13.3

Notes: Data on paypoint locations obtained from published South African National Treasury report in 2017.

Table 1.6: Summary Statistics For Municipalities - With/Without Paypoints

	At Least One Precinct Without Paypoints (1)	All precincts With Paypoints (2)	Difference (3)
Police Precinct Population	21,101	58,078	36,977**
	(15,687)	(40,085)	(11,338)
Percent Change in Municipal Population 2001-2011	1.56	-0.03	-1.59**
	(2.24)	(0.90)	(0.57)
Metropolitan Municipality	14.29	0.00	-14.29
	(36.31)	(0.00)	(8.28)
HHs in Tribal Settlements	45.94	50.97	5.03
	(35.43)	(44.21)	(14.36)
HHs in Rural Areas	48.40	53.40	5.00
	(33.68)	(42.62)	(13.78)
Mean Age	28.65	27.51	-1.14
	(1.88)	(2.81)	(0.87)
Population under 15	29.47	36.33	6.85***
	(3.73)	(5.10)	(1.61)
Population between 15 and 64	63.24	56.39	-6.85***
	(4.49)	(4.43)	(1.57)
Population over 64	7.27	7.26	-0.01
	(1.47)	(1.96)	(0.62)
No formal schooling	9.87	14.05	4.18*
	(4.41)	(5.18)	(1.72)
12 years of schooling	18.78	13.74	-5.04**
	(5.84)	(2.91)	(1.54)
Higher Education	8.09	5.32	-2.77**
	(3.11)	(1.62)	(0.83)
Avg HH Size	3.41	3.92	0.51**
	(0.18)	(0.53)	(0.15)
Female-Headed Households	44.18	53.63	9.45***
	(5.88)	(7.87)	(2.50)
Formal Dwellings	77.12	50.45	-26.67***
	(17.50)	(21.78)	(7.08)
Total Municipal Operating Expenditure	9,059	1,356	-7,703
	(18,688)	(1,104)	(4,273)
Unemployment Rate	32.09	42.46	10.37**
	(8.63)	(9.24)	(3.17)
Youth Unemp. Rate (15-34)	40.36	51.95	11.60**
	(10.23)	(10.78)	(3.72)
Percent Poor	20.13	20.21	0.08
	(3.53)	(4.44)	(1.44)
Males per 100 Females	92.54	85.79	-6.75*
	(3.98)	(8.93)	(2.56)
Percent White	4.07	4.18	0.11
	(4.40)	(5.41)	(1.76)
HH w/o piped water	47.60	53.42	5.82
	(27.84)	(36.41)	(11.65)
Other Language at Home	2.46	2.02	-0.44
	(1.70)	(1.45)	(0.55)
N	14	19	33

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

Estimates based on municipality-level characteristics measured in 2011 from the South African census. Standard errors are in parentheses. Municipalities defined as having paypoints are those where all precincts within the municipality had at least one paypoint location. Those without have at least one precinct without paypoints. Population change is measured from 2001 to 2011. The government defines two municipalities as metropolitan areas, Nelson Mandela Bay and Buffalo City municipalities. The sex ratio is defined as number of males to 100 females. Finally, other language indicates whether a language outside of South Africa's 11 official languages are spoken as main language at home.

Table 1.7: Summary Statistics For Municipalities - High/Low Paypoints

	All Precincts Low No. Paypoints	At Least One Precinct High No. Paypoints	Diff.	At Least One Precinct Low No. Paypoints	All Precincts High No. Paypoints	Diff.
	(1)	(2)	(3)	(4)	(5)	(6)
Police Precinct Population	15,909	47,120	31,210	25,046	77,081	52,035***
	(9,452)	(37,836)	(17,223)	(19,345)	(39,279)	(10,121)
% Δ in Muni. Pop 2001-2011	2.42	0.32	-2.10*	1.03	-0.14	-1.17
	(3.21)	(1.22)	(0.79)	(2.00)	(0.76)	(0.63)
Metropolitan Municipality	0.00	7.14	7.14	9.09	0.00	-9.09
	(0.00)	(26.23)	(11.88)	(29.42)	(0.00)	(8.94)
HHs in Tribal Settlements	45.06	49.51	4.45	49.67	47.15	-2.52
	(41.57)	(40.69)	(19.81)	(37.85)	(46.45)	(15.07)
HHs in Rural Areas	47.11	52.02	4.91	52.01	49.80	-2.20
	(40.92)	(38.90)	(19.01)	(36.56)	(44.20)	(14.47)
Mean Age	28.14	27.97	-0.18	28.24	27.50	-0.74
	(1.86)	(2.61)	(1.23)	(2.09)	(3.20)	(0.92)
Population under 15	26.92	34.58	7.66**	30.78	38.69	7.91***
	(2.20)	(5.30)	(2.43)	(4.60)	(3.53)	(1.58)
Population between 15 and 64	66.12	58.08	-8.05**	62.00	53.89	-8.10***
	(2.44)	(5.08)	(2.34)	(4.78)	(1.81)	(1.50)
Population over 64	6.90	7.33	0.43	7.21	7.38	0.17
	(1.79)	(1.76)	(0.86)	(1.64)	(2.01)	(0.65)
No formal schooling	9.20	12.82	3.63	10.70	15.43	4.73*
	(1.79)	(5.47)	(2.50)	(4.67)	(5.03)	(1.77)
12 years of schooling	19.38	15.25	-4.13	17.77	12.09	-5.68**
	(2.72)	(5.09)	(2.35)	(4.99)	(2.02)	(1.58)
Higher Education	7.58	6.30	-1.28	7.35	4.78	-2.56**
	(3.76)	(2.52)	(1.32)	(2.93)	(0.82)	(0.91)
Avg HH Size	3.36	3.76	0.40	3.53	4.05	0.51**
	(0.21)	(0.50)	(0.23)	(0.38)	(0.50)	(0.16)
Female-Headed Households	40.44	51.26	10.82**	45.81	57.24	11.42***
	(3.65)	(8.01)	(3.69)	(7.72)	(2.75)	(2.42)
Formal Dwellings	89.30	56.85	-32.45**	73.92	37.45	-36.48***
	(6.57)	(22.45)	(10.24)	(17.88)	(13.35)	(6.11)
Total Muni. Operating Expend.	1,670	5,151	3,481	6,387	1,097	-5,291
	(769)	(13,590)	(6,159)	(15,173)	(261)	(4,612)
Unemployment Rate	26.84	40.07	13.23**	33.54	47.12	13.58***
	(7.08)	(9.49)	(4.47)	(9.36)	(4.24)	(2.98)
Youth Unemp. Rate (15-34)	34.66	49.24	14.58**	42.13	56.85	14.72***
	(9.48)	(11.01)	(5.26)	(11.33)	(5.01)	(3.60)
Percent Poor	19.20	20.35	1.15	20.27	20.00	-0.27
	(4.71)	(3.96)	(1.97)	(3.68)	(4.82)	(1.51)
Males per 100 Females	94.90	87.54	-7.36	89.96	86.05	-3.90
	(5.59)	(7.83)	(3.68)	(9.42)	(1.82)	(2.89)
Percent White	3.49	4.25	0.77	3.99	4.43	0.44
	(4.08)	(5.12)	(2.43)	(4.94)	(5.13)	(1.85)
HH w/o piped water	44.47	52.11	7.64	51.40	50.05	-1.35
	(33.17)	(33.09)	(16.07)	(31.13)	(37.22)	(12.27)
Other Language at Home	3.35	2.00	-1.35	2.20	2.20	-0.00
	(2.57)	(1.26)	(0.72)	(1.46)	(1.79)	(0.58)
N	5	28	33	22	11	33

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

Estimates based on municipality-level characteristics measured in 2011 from the South African census. Standard errors are in parentheses. Municipalities defined as having high number paypoints are those where all precincts within the municipality had above median number of paypoint locations. Population change is measured from 2001 to 2011. The government defines two municipalities as metropolitan areas, Nelson Mandela Bay and Buffalo City municipalities. The sex ratio is defined as number of males to 100 females. Finally, other language indicates whether a language outside of South Africa's 11 official languages are spoken as main language at home.

Table 1.8: DiD With/Without Paypoints: Test of Parallel Trends on Non-Payment Days

	All Crime (1)	Violent (2)
Day of Month x With Paypoints	-0.000	0.001
	(0.002)	(0.002)
With Paypoints	0.656*	0.236
	(0.305)	(0.161)
Day of Month	-0.005*	-0.007**
	(0.002)	(0.002)
Muni. Controls	✓	✓
Weather effects	✓	✓
Holiday effects	✓	✓
Day-of-week effects	✓	✓
Day-of-month effects	X	X
Month x Year effects	✓	✓
N	120,672	120,672

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

Standard errors clustered at the precinct level. Estimates show results of a test for differential trends on daily crime counts at the precinct level for precincts with and without paypoints on days excluding social grant payment dates. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects. Municipality-level demographic characteristics are included.

Table 1.9: DiD With/Without Paypoints: Test of Parallel Trends on Non-Payment Days - Days 8 to 21

	All Crime (1)	Violent (2)
Day of Month x With Paypoints	0.002	0.003
	(0.005)	(0.005)
With Paypoints	0.579*	0.158
	(0.271)	(0.127)
Day of Month	0.005	-0.003
	(0.004)	(0.005)
Muni. Controls	✓	✓
Weather effects	✓	✓
Holiday effects	✓	✓
Day-of-week effects	✓	✓
Day-of-month effects	X	X
Month x Year effects	✓	✓
N	56,109	56,109

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

Standard errors clustered at the precinct level. Estimates show results of a test for differential trends on daily crime counts at the precinct level for precincts with and without paypoints during the second and third weeks of the month days excluding social grant payment dates. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects. Municipality-level demographic characteristics are included.

Table 1.10: DiD With/Without Paypoints: Test of Parallel Trends on Non-Payment Days - Days 8 to 14

	All Crime (1)	Violent (2)
Day of Month x With Paypoints	0.019	-0.005
	(0.015)	(0.015)
With Paypoints	0.395	0.236
	(0.362)	(0.226)
Day of Month	-0.016	-0.010
	(0.014)	(0.014)
Muni. Controls	✓	✓
Weather effects	✓	✓
Holiday effects	✓	✓
Day-of-week effects	✓	✓
Day-of-month effects	X	X
Month x Year effects	✓	✓
N	27,778	27,778

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

Standard errors clustered at the precinct level. Estimates show results of a test for differential trends on daily crime counts at the precinct level for precincts with and without paypoints during the second week of the month days excluding social grant payment dates. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects. Municipality-level demographic characteristics are included.

Table 1.11: DiD With/Without Paypoints: Effects of Social Grant Payment Dates on Daily Crime Counts

	All Crime (1)	All Crime (2)	Violent (3)	Violent (4)
Prior Day x With Paypoints	-0.163	-0.179*	-0.065	-0.079
	(0.092)	(0.084)	(0.111)	(0.108)
With Paypoints	0.709		0.278	
	(0.386)		(0.187)	
Prior Day	0.310***	0.359***	0.262*	0.285**
	(0.084)	(0.079)	(0.105)	(0.103)
Adj. R-squared	0.130	0.491	0.116	0.334
Pay Day x With Paypoints	0.035	0.018	0.052	0.039
	(0.077)	(0.063)	(0.042)	(0.039)
With Paypoints	0.703		0.275	
	(0.387)		(0.190)	
Pay Day	0.294***	0.385***	0.133**	0.171***
	(0.080)	(0.078)	(0.041)	(0.043)
Adj. R-squared	0.131	0.492	0.116	0.334
Next Day x With Paypoints	0.220**	0.216**	0.142*	0.136*
	(0.082)	(0.076)	(0.059)	(0.055)
With Paypoints	0.697		0.272	
	(0.388)		(0.189)	
Next Day	-0.031	-0.007	0.049	0.058
	(0.065)	(0.065)	(0.050)	(0.049)
Adj. R-squared	0.130	0.491	0.116	0.334
P-values of testing coefficients of interaction terms				
Prior Day = Payday	0.092	0.087	0.367	0.361
Next Day = Payday	0.096	0.075	0.166	0.147
Mean Daily Count	2.407	2.407	1.036	1.036
Police Precinct FEs	X	✓	X	✓
Muni Controls	✓	X	✓	X
Weather effects	✓	✓	✓	✓
Holiday effects	✓	✓	✓	✓
Day-of-week effects	✓	✓	✓	✓
Day-of-month effects	✓	✓	✓	✓
Month x Year effects	✓	✓	✓	✓
N	125,214	125,214	125,214	125,214

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

Notes: Estimates based on data from April 2014 to September 2018, before the mass decommissioning of paypoint locations. Estimates present standard errors clustered at the police precinct level. Sample includes crimes for which SAPS provided coordinate locations or place or neighborhood information. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects and month by year effects. Quadratic day of month control and municipality controls are also included.

Table 1.12: DiD High/Low Paypoints: Test of Parallel Trends on Non-Payment Days

	All Crime (1)	Violent (2)
Day of Month x High No. Paypoints	0.002	0.003
	(0.002)	(0.002)
High No. Paypoints	-0.459	-0.238
	(0.483)	(0.227)
Day of Month	-0.006***	-0.008***
	(0.002)	(0.001)
Muni. Controls	✓	✓
Weather effects	✓	✓
Holiday effects	✓	✓
Day-of-week effects	✓	✓
Day-of-month effects	X	X
Month x Year effects	✓	✓
N	119,227	119,227

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

Standard errors clustered at the precinct level. Estimates show results of a test for differential trends on daily crime counts at the precinct level for precincts with above and below median number of paypoint locations per 10,000 residents on days excluding social grant payment dates. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects. Municipality-level demographic characteristics are included.

Table 1.13: DiD High/Low Paypoints: Test of Parallel Trends on Non-Payment Days - Days 8 to 21

	All Crime (1)	Violent (2)
Day of Month x High No. Paypoints	-0.000	0.002
	(0.004)	(0.003)
High No. Paypoints	-0.403	-0.201
	(0.444)	(0.189)
Day of Month	0.008*	-0.001
	(0.003)	(0.002)
Muni. Controls	✓	✓
Weather effects	✓	✓
Holiday effects	✓	✓
Day-of-week effects	✓	✓
Day-of-month effects	X	X
Month x Year effects	✓	✓
N	55,422	55,422

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

Standard errors clustered at the precinct level. Estimates show results of a test for differential trends on daily crime counts at the precinct level for precincts with above and below median number of paypoint locations per 10,000 residents during the second and third weeks of the month excluding social grant payment dates. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects. Municipality-level demographic characteristics are included.

Table 1.14: DiD High/Low Paypoints: Test of Parallel Trends on Non-Payment Days - Days 8 to 14

	All Crime (1)	Violent (2)
Day of Month x High No. Paypoints	-0.002	-0.001
	(0.011)	(0.009)
High No. Paypoints	-0.392	-0.184
	(0.473)	(0.224)
Day of Month	0.001	-0.015
	(0.008)	(0.008)
Muni. Controls	✓	✓
Weather effects	✓	✓
Holiday effects	✓	✓
Day-of-week effects	✓	✓
Day-of-month effects	X	X
Month x Year effects	✓	✓
N	27,438	27,438

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

Standard errors clustered at the precinct level. Estimates show results of a test for differential trends on daily crime counts at the precinct level for precincts with above and below median number of paypoint locations per 10,000 residents during the second week of the month excluding social grant payment dates. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects. Municipality-level demographic characteristics are included.

Table 1.15: Generalized DiD: Test of Parallel Trends on Non-Payment Days - Days 8 to 21

	All Crime (1)	Violent (2)
Day of Month x No. Paypoints	0.000	0.000
	(0.000)	(0.000)
No. Paypoints	0.008	0.006
	(0.011)	(0.004)
Day of Month	0.048***	-0.014
	(0.013)	(0.011)
Muni. Controls	✓	✓
Weather effects	✓	✓
Holiday effects	✓	✓
Day-of-week effects	✓	✓
Day-of-month effects	X	X
Month x Year effects	✓	✓
N	55,422	55,422

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

Standard errors clustered at the precinct level. Estimates show results of a test for differential trends on daily crime counts at the precinct level for precincts with different numbers of paypoint locations per 10,000 residents on days excluding social grant payment dates. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects. Municipality-level demographic characteristics are included.

Table 1.16: DiD High/Low Paypoints: Effects of Social Grant Payment Dates on Daily Crime Counts

	All Crime (1)	All Crime (2)	Violent (3)	Violent (4)
Prior Day x High No. Paypoints	-0.151*	-0.152*	-0.245***	-0.243***
	(0.067)	(0.066)	(0.060)	(0.059)
High No. Paypoints	-0.421		-0.187	
	(0.468)		(0.208)	
Prior Day	0.242***	0.262***	0.320***	0.323***
	(0.054)	(0.056)	(0.049)	(0.049)
Adj. R-squared	0.171	0.492	0.143	0.335
Pay Day x High No. Paypoints	0.191**	0.207***	0.194***	0.205***
	(0.060)	(0.057)	(0.042)	(0.043)
High No. Paypoints	-0.433		-0.202	
	(0.468)		(0.209)	
Pay Day	0.255***	0.311***	0.099***	0.115***
	(0.052)	(0.058)	(0.029)	(0.032)
Adj. R-squared	0.172	0.493	0.143	0.335
Next Day x High No. Paypoints	-0.016	-0.010	0.044	0.051
	(0.077)	(0.072)	(0.050)	(0.048)
High No. Paypoints	-0.425		-0.197	
	(0.467)		(0.208)	
Next Day	0.188**	0.200**	0.164***	0.162***
	(0.068)	(0.067)	(0.042)	(0.042)
Adj. R-squared	0.171	0.492	0.142	0.335
P-values of testing coefficients of interaction terms				
Prior Day = Payday	0.000	0.000	0.000	0.000
Next Day = Payday	0.007	0.007	0.013	0.013
Mean Daily Count	3.062	3.062	1.277	1.277
Police Precinct FEs	X	✓	X	✓
Muni. Controls	✓	X	✓	X
Weather effects	✓	✓	✓	✓
Holiday effects	✓	✓	✓	✓
Day-of-week effects	✓	✓	✓	✓
Day-of-month effects	✓	✓	✓	✓
Month x Year effects	✓	✓	✓	✓
N	123,714	123,714	123,714	123,714

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

Notes: Estimates based on data from April 2014 to September 2018, before the mass decommissioning of paypoint locations. Precincts with a high number of paypoints are defined as those with more than the median number of paypoints per 10,000 residents. Estimates present standard errors clustered at the police precinct level. Sample includes crimes for which SAPS provided coordinate locations or place or neighborhood information. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects and month by year effects. Quadratic day of month control and municipality controls are also included. Mean daily count is the mean for precincts with lower relative number of paypoint locations.

Table 1.17: Generalized DiD: Effects of Social Grant Payment Dates on Daily Crime Counts

	All Crime (1)	All Crime (2)	Violent (3)	Violent (4)
Prior Day x No. Paypoints	-0.001	-0.002	-0.003	-0.003
	(0.006)	(0.006)	(0.009)	(0.009)
No. Paypoints	0.013		0.010	
	(0.012)		(0.005)	
Prior Day	0.179***	0.201***	0.219***	0.224***
	(0.042)	(0.044)	(0.044)	(0.045)
Adj. R-squared	0.180	0.492	0.147	0.335
Pay Day x No. Paypoints	0.003	0.003	0.004	0.003
	(0.006)	(0.006)	(0.006)	(0.006)
No. Paypoints	0.013		0.010	
	(0.012)		(0.006)	
Pay Day	0.328***	0.393***	0.171***	0.193***
	(0.053)	(0.062)	(0.035)	(0.039)
Adj. R-squared	0.180	0.493	0.147	0.335
Next Day x No. Paypoints	-0.001	-0.002	0.002	0.002
	(0.002)	(0.002)	(0.001)	(0.002)
No. Paypoints	0.013		0.010	
	(0.012)		(0.006)	
Next Day	0.184***	0.204***	0.174***	0.177***
	(0.048)	(0.051)	(0.031)	(0.033)
Adj. R-squared	0.180	0.492	0.147	0.335
P-values of testing coefficients of interaction terms				
Prior Day = Payday	0.712	0.697	0.666	0.661
Next Day = Payday	0.485	0.464	0.834	0.816
Mean Daily Count	3.062	3.062	1.277	1.277
Police Precinct FEs	X	✓	X	✓
Precinct Pop	✓	X	✓	X
Weather effects	✓	✓	✓	✓
Holiday effects	✓	✓	✓	✓
Day-of-week effects	✓	✓	✓	✓
Day-of-month effects	✓	✓	✓	✓
Month x Year effects	✓	✓	✓	✓
N	123,714	123,714	123,714	123,714

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

Notes: Estimates based on data from April 2014 to September 2018, before the mass decommissioning of paypoint locations. Paypoints are measured per 10,000 residents and sample excludes precincts with zero paypoint locations. Estimates present standard errors clustered at the police precinct level. Sample includes crimes for which SAPS provided coordinate locations or place or neighborhood information. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects and month by year effects. Quadratic day of month control is also included. Mean daily count is the mean for precincts with below the median number of paypoint locations.

Table 1.18: DiD With/Without Paypoints: Effects of Social Grant Payment Dates on Weekly Crime Counts

	All Crime (1)	All Crime (2)	Violent (3)	Violent (4)
SG Week x With Paypoints	-0.115	-0.182	-0.023	-0.055
	(0.509)	(0.510)	(0.149)	(0.148)
With Paypoints	2.185		0.824	
	(1.215)		(0.552)	
SG Week	-1.644***	-1.670***	-0.523***	-0.534***
	(0.478)	(0.474)	(0.137)	(0.135)
Adj. R-squared	0.231	0.594	0.175	0.521
Mean Weekly Count	6.857	6.857	2.959	2.959
Police Precinct FEs	X	✓	X	✓
Muni Controls	✓	X	✓	X
Weather effects	✓	✓	✓	✓
Month x Year effects	✓	✓	✓	✓
N	46,587	46,587	46,587	46,587

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

Notes: Estimates based on data from April 2014 to September 2018, before the mass decommissioning of paypoint locations. Outcome variables are measured at weekly count levels. Estimates present standard errors clustered at the police precinct level. Sample includes crimes for which SAPS provided coordinate locations or place or neighborhood information. Weather effects include weekly average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Time fixed effects include month by year effects. Municipality controls are also included. Mean weekly count is the mean for precincts without paypoint locations

Table 1.19: DiD High/Low Paypoints: Effects of Social Grant Payment Dates on Weekly Crime Counts

	All Crime (1)	All Crime (2)	Violent (3)	Violent (4)
SG Week x High No. Paypoints	1.093***	1.189***	0.336**	0.385***
	(0.321)	(0.347)	(0.105)	(0.113)
High No. Paypoints	-2.050		-0.820	
	(1.635)		(0.724)	
SG Week	-2.288***	-2.418***	-0.708***	-0.770***
	(0.275)	(0.284)	(0.088)	(0.092)
Adj. R-squared	0.232	0.595	0.177	0.522
Mean Weekly Count	9.092	9.092	3.792	3.792
Police Precinct FEs	X	✓	X	✓
Muni Controls	✓	X	✓	X
Weather effects	✓	✓	✓	✓
Month x Year effects	✓	✓	✓	✓
N	46,059	46,059	46,059	46,059

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

Notes: Estimates based on data from April 2014 to September 2018, before the mass decommissioning of paypoint locations. Outcome variables are measured at weekly count levels. Precincts with a high number of paypoints are defined as those with more than the median number of paypoints per 10,000 residents. Estimates present standard errors clustered at the police precinct level. Sample includes crimes for which SAPS provided coordinate locations or place or neighborhood information. Weather effects include weekly average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Time fixed effects include month by year effects. Municipality controls are also included. Mean weekly count is the mean for precincts without paypoint locations

Table 1.20: DiD With/Without Paypoints: Excluding Largest Two Cities

	All Crime (1)	All Crime (2)	Violent (3)	Violent (4)
Prior Day x With Paypoints	-0.142	-0.173	-0.084	-0.109
	(0.131)	(0.114)	(0.074)	(0.062)
With Paypoints	0.806		0.428	
	(0.494)		(0.260)	
Prior Day	0.282*	0.336**	0.249***	0.278***
	(0.128)	(0.114)	(0.068)	(0.058)
Adj. R-squared	0.127	0.475	0.123	0.321
Pay Day x With Paypoints	0.151	0.132	0.042	0.026
	(0.082)	(0.072)	(0.058)	(0.060)
With Paypoints	0.798		0.424	
	(0.495)		(0.261)	
Pay Day	0.153	0.227**	0.136*	0.172**
	(0.080)	(0.083)	(0.057)	(0.065)
Adj. R-squared	0.128	0.476	0.123	0.321
Next Day x With Paypoints	0.256*	0.261*	0.119	0.116
	(0.105)	(0.104)	(0.107)	(0.104)
With Paypoints	0.794		0.422	
	(0.494)		(0.260)	
Next Day	-0.080	-0.079	0.058	0.059
	(0.086)	(0.091)	(0.101)	(0.099)
Adj. R-squared	0.127	0.475	0.123	0.321
P-values of testing coefficients of interaction terms				
Prior Day = Payday	0.097	0.080	0.256	0.227
Next Day = Payday	0.474	0.391	0.582	0.523
Mean Daily Count	1.603	1.603	0.790	0.790
Police Precinct FEs	X	✓	X	✓
Muni Controls	✓	X	✓	X
Weather effects	✓	✓	✓	✓
Holiday effects	✓	✓	✓	✓
Day-of-week effects	✓	✓	✓	✓
Day-of-month effects	✓	✓	✓	✓
Month x Year effects	✓	✓	✓	✓
N	93,839	93,839	93,839	93,839

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

Notes: Estimates based on data from April 2014 to September 2018, before the mass decommissioning of paypoint locations. Estimates present standard errors clustered at the police precinct level. Sample includes crimes for which SAPS provided coordinate locations or place or neighborhood information. The two largest cities in the Eastern Cape province, Port Elizabeth and East London are excluded from the sample. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects and month by year effects. Quadratic day of month control and municipality controls are also included. Mean daily count is the mean for precincts without paypoint locations

Table 1.21: DiD High/Low Paypoints: Excluding Largest Two Cities

	All Crime (1)	All Crime (2)	Violent (3)	Violent (4)
Prior Day x High No. Paypoints	-0.103	-0.093	-0.165**	-0.155**
	(0.059)	(0.055)	(0.060)	(0.057)
High No. Paypoints	-0.380		-0.101	
	(0.496)		(0.235)	
Prior Day	0.214***	0.224***	0.267***	0.264***
	(0.050)	(0.050)	(0.051)	(0.050)
Adj. R-squared	0.126	0.475	0.121	0.321
Pay Day x High No. Paypoints	0.159*	0.189**	0.138**	0.156**
	(0.075)	(0.070)	(0.051)	(0.052)
High No. Paypoints	-0.389		-0.112	
	(0.495)		(0.236)	
Pay Day	0.209**	0.246**	0.097*	0.106*
	(0.067)	(0.073)	(0.042)	(0.045)
Adj. R-squared	0.127	0.476	0.121	0.321
Next Day x High No. Paypoints	-0.026	-0.003	0.018	0.034
	(0.104)	(0.094)	(0.065)	(0.060)
High No. Paypoints	-0.383		-0.107	
	(0.494)		(0.234)	
Next Day	0.185	0.173	0.165**	0.151**
	(0.096)	(0.090)	(0.058)	(0.054)
Adj. R-squared	0.126	0.475	0.121	0.321
P-values of testing coefficients of interaction terms				
Prior Day = Payday	0.004	0.002	0.000	0.000
Next Day = Payday	0.047	0.050	0.078	0.080
Mean Daily Count	2.424	2.424	1.096	1.096
Police Precinct FEs	X	✓	X	✓
Muni. Controls	✓	X	✓	X
Weather effects	✓	✓	✓	✓
Holiday effects	✓	✓	✓	✓
Day-of-week effects	✓	✓	✓	✓
Day-of-month effects	✓	✓	✓	✓
Month x Year effects	✓	✓	✓	✓
N	93,564	93,564	93,564	93,564

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

Notes: Estimates based on data from April 2014 to September 2018, before the mass decommissioning of paypoint locations. Precincts with a high number of paypoints are defined as those with more than the median number of paypoints per 10,000 residents. Estimates present standard errors clustered at the police precinct level. Sample includes crimes for which SAPS provided coordinate locations or place or neighborhood information. The two largest cities in the Eastern Cape province, Port Elizabeth and East London are excluded from the sample. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects and month by year effects. Quadratic day of month control and municipality controls are also included. Mean daily count is the mean for precincts lower relative number of paypoint locations.

Table 1.22: DiD High/Low Paypoints: Effects of Social Grant Payment Dates on Daily Crime Counts - Excluding Police Precincts with Zero Paypoint Locations

	All Crime (1)	All Crime (2)	Violent (3)	Violent (4)
Prior Day x High No. Paypoints	-0.085	-0.079	-0.214***	-0.208***
	(0.070)	(0.068)	(0.063)	(0.061)
High No. Paypoints	-0.496		-0.208	
	(0.451)		(0.198)	
Prior Day	0.163**	0.174**	0.349***	0.347***
	(0.062)	(0.062)	(0.056)	(0.055)
Adj. R-squared	0.142	0.455	0.077	0.264
Pay Day x High No. Paypoints	0.186**	0.199**	0.190***	0.201***
	(0.067)	(0.061)	(0.046)	(0.045)
High No. Paypoints	-0.505		-0.222	
	(0.450)		(0.199)	
Pay Day	-0.096*	-0.101*	-0.202***	-0.213***
	(0.045)	(0.041)	(0.032)	(0.032)
Adj. R-squared	0.142	0.455	0.076	0.264
Next Day x High No. Paypoints	-0.028	-0.019	0.045	0.054
	(0.082)	(0.072)	(0.051)	(0.047)
High No. Paypoints	-0.497		-0.217	
	(0.450)		(0.198)	
Next Day	0.131	0.129*	0.081	0.072
	(0.071)	(0.064)	(0.043)	(0.040)
Adj. R-squared	0.142	0.455	0.076	0.263
P-values of testing coefficients of interaction terms				
Prior Day = Payday	0.008	0.006	0.000	0.000
Next Day = Payday	0.005	0.006	0.014	0.014
Mean Daily Count	3.168	3.168	1.316	1.316
Police Precinct FEs	X	✓	X	✓
Muni. Controls	✓	X	✓	X
Weather effects	✓	✓	✓	✓
Holiday effects	✓	✓	✓	✓
Day-of-week effects	✓	✓	✓	✓
Day-of-month effects	✓	✓	✓	✓
Month x Year effects	✓	✓	✓	✓
N	114,198	114,198	114,198	114,198

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

Notes: Estimates based on data from April 2014 to September 2018, before the mass decommissioning of paypoint locations. Precincts with a high number of paypoints are defined as those with more than the median number of paypoints per 10,000 residents. Estimates present standard errors clustered at the police precinct level. Sample includes crimes for which SAPS provided coordinate locations or place or neighborhood information. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects and month by year effects. Quadratic day of month control is also included. Municipality controls are also included. Mean daily count is the mean for precincts with lower relative number of paypoint locations. Sample excludes precincts with zero paypoints.

Table 1.23: Generalized DiD: Effects of Social Grant Payment Dates on Daily Crime Counts - Excluding Police Precincts with Zero Paypoint Locations

	All Crime (1)	All Crime (2)	Violent (3)	Violent (4)
Prior Day x No. Paypoints	-0.000	-0.001	-0.002	-0.002
	(0.006)	(0.006)	(0.009)	(0.009)
No. Paypoints	0.009		0.009	
	(0.014)		(0.006)	
Prior Day	0.167***	0.192***	0.216***	0.221***
	(0.046)	(0.048)	(0.047)	(0.048)
Adj. R-squared	0.181	0.498	0.146	0.340
Pay Day x No. Paypoints	0.003	0.003	0.003	0.003
	(0.007)	(0.006)	(0.006)	(0.006)
No. Paypoints	0.009		0.009	
	(0.014)		(0.007)	
Pay Day	0.324***	0.393***	0.171***	0.194***
	(0.057)	(0.066)	(0.038)	(0.043)
Adj. R-squared	0.181	0.498	0.146	0.340
Next Day x No. Paypoints	-0.002	-0.003	0.002	0.001
	(0.002)	(0.003)	(0.002)	(0.002)
No. Paypoints	0.009		0.009	
	(0.014)		(0.006)	
Next Day	0.199***	0.223***	0.182***	0.186***
	(0.053)	(0.055)	(0.034)	(0.035)
Adj. R-squared	0.181	0.498	0.146	0.340
P-values of testing coefficients of interaction terms				
Prior Day = Payday	0.770	0.754	0.729	0.689
Next Day = Payday	0.426	0.410	0.693	0.780
Mean Daily Count	3.168	3.168	1.316	1.316
Police Precinct FEs	X	✓	X	✓
Precinct Pop	✓	X	✓	X
Weather effects	✓	✓	✓	✓
Holiday effects	✓	✓	✓	✓
Day-of-week effects	✓	✓	✓	✓
Day-of-month effects	✓	✓	✓	✓
Month x Year effects	✓	✓	✓	✓
N	114,198	114,198	114,198	114,198

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

Notes: Estimates based on data from April 2014 to September 2018, before the mass decommissioning of paypoint locations. Paypoints are measured per 10,000 residents and sample excludes precincts with zero paypoint locations. Estimates present standard errors clustered at the police precinct level. Sample includes crimes for which SAPS provided coordinate locations or place or neighborhood information. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects and month by year effects. Quadratic day of month control is also included. Mean daily count is the mean for precincts with below the median number of paypoint locations.

Table 1.24: DiD High/Low Paypoints: Effects of Social Grant Payment Dates on Daily Crime Counts - Excluding Police Precincts with Top 5% Number of Paypoint Locations

	All Crime (1)	All Crime (2)	Violent (3)	Violent (4)
Prior Day x High No. Paypoints	-0.112	-0.101	-0.219***	-0.212***
	(0.064)	(0.062)	(0.056)	(0.055)
High No. Paypoints	-0.558		-0.266	
	(0.502)		(0.215)	
Prior Day	0.176**	0.183***	0.341***	0.338***
	(0.054)	(0.055)	(0.050)	(0.050)
Adj. R-squared	0.143	0.449	0.077	0.258
Pay Day x High No. Paypoints	0.193**	0.210***	0.198***	0.209***
	(0.062)	(0.059)	(0.042)	(0.043)
High No. Paypoints	-0.569		-0.280	
	(0.502)		(0.215)	
Pay Day	-0.091*	-0.095*	-0.199***	-0.207***
	(0.038)	(0.037)	(0.028)	(0.028)
Adj. R-squared	0.143	0.449	0.076	0.257
Next Day x High No. Paypoints	0.007	0.020	0.062	0.071
	(0.074)	(0.067)	(0.046)	(0.044)
High No. Paypoints	-0.562		-0.275	
	(0.501)		(0.215)	
Next Day	0.099	0.097	0.063	0.056
	(0.061)	(0.057)	(0.037)	(0.035)
Adj. R-squared	0.143	0.449	0.076	0.257
P-values of testing coefficients of interaction terms				
Prior Day = Payday	0.001	0.001	0.000	0.000
Next Day = Payday	0.012	0.013	0.014	0.013
Mean Daily Count	3.062	3.062	1.277	1.277
Police Precinct FEs	X	✓	X	✓
Precinct Pop	✓	X	✓	X
Weather effects	✓	✓	✓	✓
Holiday effects	✓	✓	✓	✓
Day-of-week effects	✓	✓	✓	✓
Day-of-month effects	✓	✓	✓	✓
Month x Year effects	✓	✓	✓	✓
N	120,193	120,193	120,193	120,193

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

Notes: Estimates based on data from April 2014 to September 2018, before the mass decommissioning of paypoint locations. Precincts with a high number of paypoints are defined as those with more than the median number of paypoints per 10,000 residents. Estimates present standard errors clustered at the police precinct level. Sample includes crimes for which SAPS provided coordinate locations or place or neighborhood information. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays as well as the monthly salary payment dates. Time fixed effects include day of week effects and month by year effects. Quadratic day of month control is also included. Municipality controls are also included. Mean daily count is the mean for precincts with lower relative number of paypoint locations. Sample excludes precincts with more than 15.04 paypoints per 10,000 residents.

Table 1.25: DiD With/Without Paypoints: Effects of Salary Payments on Daily Crime Counts

	All Crime (1)	All Crime (2)	Violent (3)	Violent (4)
Prior Day x With Paypoints	0.080*	0.093**	0.029	0.035*
	(0.033)	(0.029)	(0.017)	(0.016)
With Paypoints	0.703		0.277	
	(0.387)		(0.189)	
Prior Day	-0.155***	-0.125***	-0.109***	-0.096***
	(0.035)	(0.028)	(0.023)	(0.021)
Adj. R-squared	0.131	0.493	0.118	0.337
Salary Payday x With Paypoints	0.170*	0.180*	0.106	0.122
	(0.083)	(0.071)	(0.090)	(0.093)
With Paypoints	0.700		0.274	
	(0.388)		(0.191)	
Salary Payday	-0.082	-0.081	-0.071	-0.082
	(0.069)	(0.060)	(0.083)	(0.086)
Adj. R-squared	0.131	0.493	0.118	0.337
Next Day x With Paypoints	0.102**	0.109***	0.046**	0.049**
	(0.032)	(0.028)	(0.016)	(0.016)
With Paypoints	0.701		0.276	
	(0.387)		(0.189)	
Next Day	0.214***	0.139***	0.164***	0.132***
	(0.047)	(0.039)	(0.027)	(0.024)
Adj. R-squared	0.131	0.493	0.118	0.337
P-values of testing coefficients of interaction terms				
Prior Day = Salary Payday	0.184	0.136	0.350	0.309
Next Day = Salary Payday	0.331	0.234	0.467	0.396
Mean Daily Count	2.407	2.407	1.036	1.036
Police Precinct FEs	X	✓	X	✓
Muni. Controls	✓	X	✓	X
Weather effects	✓	✓	✓	✓
Holiday effects	✓	✓	✓	✓
Day-of-week effects	✓	✓	✓	✓
Day-of-month effects	✓	✓	✓	✓
Month x Year effects	✓	✓	✓	✓
N	125,214	125,214	125,214	125,214

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

Notes: Estimates based on data from April 2014 to September 2018, before the mass decommissioning of paypoint locations. Estimates present standard errors clustered at the police precinct level. Sample includes crimes for which SAPS provided coordinate locations or place or neighborhood information. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate federal holidays. Time fixed effects include day of week effects and month by year effects. Quadratic day of month control, social grant payment day indicator, and municipality controls are also included.

Table 1.26: DiD High/Low Paypoints: Effects of Salary Paydays on Daily Crime Counts

	All Crime (1)	All Crime (2)	Violent (3)	Violent (4)
Prior Day x High No. Paypoints	0.119*	0.139*	0.101**	0.114**
	(0.060)	(0.057)	(0.039)	(0.039)
High No. Paypoints	-0.430		-0.199	
	(0.469)		(0.210)	
Prior Day	-0.159***	-0.127***	-0.108***	-0.094***
	(0.034)	(0.028)	(0.023)	(0.021)
Adj. R-squared	0.172	0.494	0.145	0.338
Salary Payday x High No. Paypoints	0.101	0.107	0.148*	0.158**
	(0.089)	(0.087)	(0.057)	(0.058)
High No. Paypoints	-0.429		-0.200	
	(0.469)		(0.211)	
Salary Payday	0.027	0.040	-0.042	-0.041
	(0.047)	(0.047)	(0.027)	(0.028)
Adj. R-squared	0.172	0.494	0.145	0.338
Next Day x High No. Paypoints	0.140*	0.155**	0.117**	0.128**
	(0.060)	(0.057)	(0.039)	(0.039)
High No. Paypoints	-0.430		-0.199	
	(0.469)		(0.210)	
Next Day	0.197***	0.143***	0.157***	0.135***
	(0.046)	(0.039)	(0.028)	(0.025)
Adj. R-squared	0.172	0.494	0.145	0.338
P-values of testing coefficients of interaction terms				
Prior Day = Salary Payday	0.697	0.480	0.079	0.096
Next Day = Salary Payday	0.387	0.289	0.239	0.247
Mean Daily Count	3.062	3.062	1.277	1.277
Police Precinct FEs	X	✓	X	✓
Muni. Controls	✓	X	✓	X
Weather effects	✓	✓	✓	✓
Holiday effects	✓	✓	✓	✓
Day-of-week effects	✓	✓	✓	✓
Day-of-month effects	✓	✓	✓	✓
Month x Year effects	✓	✓	✓	✓
N	123,714	123,714	123,714	123,714

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

Notes: Estimates based on data from April 2014 to September 2018, before the mass decommissioning of paypoint locations. Precincts with a high number of paypoints are defined as those with more than the median number of paypoints per 10,000 residents. Estimates present standard errors clustered at the police precinct level. Sample includes crimes for which SAPS provided coordinate locations or place or neighborhood information. Weather effects include daily average precipitation and maximum temperature measured by weather stations in the Eastern Cape. Holiday effects indicate days that are national holidays. Time fixed effects include day of week effects and month by year effects. Quadratic day of month control, social grant payday indicator, and municipality controls are also included. Finally, mean daily count is the mean for precincts with lower relative number of paypoint locations.

Chapter 2: The “Dr.” Effect

2.1 Introduction

In 2021, women represented 46 percent of all doctoral graduates and over 50 percent of graduates in fields such as biological and biomedical sciences, health sciences, and psychology [64]. The trend in women’s share of doctoral degrees has grown over the past two decades, despite a decrease in number of total graduates following the COVID-19 pandemic. Although women are not attaining doctoral training at lower rates than their male colleagues, they are still less likely to be referred to with their earned “Dr.” title in academic and professional settings [65–67]. This phenomenon played out very publicly in the case of Dr. Jill Biden, the first lady of the Biden-Harris administration who earned her Ph.D. in Education, and many prominent media outlets took to their op-ed sections to critique or express their support for the use of the honorific [68–70]. This controversy arose shortly after the findings that in academic settings, men are more likely than women to be referred to by their last names and that using a person’s last name signifies distinction [71]. Moreover, a recent study by the Gender Institute for Global Women’s Leadership, find women are represented significantly less in the news media as source experts, with some categories of news articles including female sources only 14.4 percent of the time compared to men [72]. Consequently, the common practice in the media of reserving the “Dr.” title solely for practicing medical doctors may reinforce and even aggravate the underlying gender bias in the perceived credibility and status of experts by removing valuable information about their level of expertise.

Given existing gender bias in perceptions of the credibility and status of experts, does the absence of professional titles in major news sources that removes information on their level of training impose a credibility penalty for women different to men? Our research aims to address this question, and we hypothesize that professional titles serve as an information signal to audiences that can either confirm or update their assumptions on the knowledge of the referenced professional. While not universally adhered to, the AP Stylebook is considered to be the standard for most print and online publications of newspapers and magazines, as well as some broadcasters and public relations firms [73]. We conduct a survey experiment on university students that asks them to rate the credibility of experts cited in major news articles when they are referenced with or without professional titles. While we are able to detect a positive credibility effect of using professional titles, we are unable to distinguish a differential credibility impact across gender.

2.2 Literature Review

There exists a limited but growing literature on the impact of honorific titles such as “Dr.” on the perceived credibility of researchers. For example, Atir and Ferguson [71] focus on whether the cited experts are perceived as “eminent”, defined as “fame or recognized superiority, especially within a particular sphere or profession”, “better known,” and/or “more distinguished.” While setting aside the issue of gender for this particular study, Atir and Ferguson [71] find that after having participants read and rank the biographies of researchers that are identical except for the use of “Dr.” in the treatment group, those with the extra title are perceived to have a higher degree of these qualities. In a subsequent study in the same paper, Atir and Ferguson [71] find that the perception of credibility carries over to students' perceptions of who should win major career advancements such as the National Science Foundation career award. Long. et al. [74] further develop the impact of the usage of different titles by noting that for outreach to policymakers on substances use disorders, using Ph.D. in the subject line of the email resulted in more clicks and opens of the email and content than simply using the title "Dr.", and both were more effective than the use of no title. These studies, among others, offer evidence on the measured effects the use of professional titles may have on perceptions of credibility.

Beyond the gender credibility gap in authorship, evidence shows that gender bias extends to academia through the relationship between students and professors, among other avenues. Using Amazon M-Turk for a series of seven studies, Atir and Ferguson [71] show that men are both referred to by surname almost twice as much as women and receive as much as a 14 percent increase in perceived credibility from the use of a surname. This indicates that there not only exists a difference in how women are referred to in academic settings, but also that the difference is a result of a perception of lower status. The fourth study of Atir and Ferguson [71] asked participants to write a profile of a researcher, and found that when the gender was male, respondents would refer to them by surname almost four times more than they would for females, who they referred to with first names. Moreover, several studies spanning many disciplines have found that women are cited less than men in academic journals [75–78]. For example, Chaterjee and Werner examine 5,554 articles and find that articles with female authors as primary or senior authors are cited about half as much as those with male primary or senior authors [75]. Moreover, work by King, Bergsrom and West find that women are less likely to cite their own previous work compared to their male colleagues [79]. Though this strays from our central focus on the use of titles, it offers insight into a potential credibility hierarchy: use of doctoral title on top, surname second, and first name at the bottom.

Much of the existing literature on the credibility gender gap focuses on direct authorship of material, independent of educational attainment, unlike our present study. Paul, Sui and Searles [80] focus on source credibility, specifically in the field of journalism authorship. They find little relation between the gender of the author of a story and the perceived source credibility. However, Armstrong and McAdams [81] find evidence of a gender credibility gap in a similar setting. After having participants rate the credibility of randomized blog posts written by a male or female, results suggest that blog posts written by men were perceived as more credible than those by women. When asked to guess the gender of the experts cited in the blog posts, respondents more often attributed the experts' gender to male. This finding is extended by Armstrong and Nelson [82], who find that readers assume the gender of a cited expert is male unless told otherwise. Building on Armstrong and McAdams' [81] findings, Klaas and Boukes [83] found that attributing a news article to a male journalist increased the credibility rating, even more so when the article was about something socially perceived as masculine. They further find that younger males rated female writers as less credible on average when compared to older males' ratings. These studies collectively validate the concern that women are typically found less credible as source experts by audiences.

The gender gap in authorship credibility extends to the perceptions within the academia, as MacNell, Driscoll and Hunt [66] present evidence that students give lower ratings to female teachers compared to their male counterparts. The authors conduct an experiment where assistant instructors for an online course operate under different genders, and subsequently find that students rated male assistant instructors higher. Given the role these ratings play in career advancement, the gender bias students exhibit in their perceptions of female instructor's effectiveness and knowledge are important dimensions when considering the components of perceptions in credibility. Bias in academia is not limited to student-instructor relationships. Peer academics also engage in such behavior, according to Wu [84] who notes the presence of demeaning and sexist language on a popular economics job forum. While similar research is lacking in other disciplines, articles have examined biases in everything from the writing of letters of recommendation to grant application processes in hard and medical sciences [85–87].

The findings of Wu, Llorens et al., Morgan, Hawkins and Lundine, and Khan et al. [84–87] link to a larger phenomenon in academia in particular: the enforcement of out-group biases against women. Work by Wu [84] cites offensive and derogatory language used both against women and subgroups of males (e.g. "homo") that can be seen as a territorial marking of one's field against outsiders (in this case, women and members of the LGBTQ+ community). Offensive language like this helps solidify an "in-group" that is defined as those who have traditionally dominated economics and wider academic circles and helps protect against the perceived threat of accomplished women [84]. Rudman [88] noted this in-group protectionism resulting in what she describes as a "backlash" against women who engaged in self-promotion. While Rudman [88] notes that the competence score of evaluated candidates in the study did not decrease for self-promoting women, their overall "hireability" was, findings that are confirmed by Rudman [89]. Fisher, Stinson and Kalajdzic [90] provide another example of this bias against women with high achievement by examining reviews of students left on RateMyProfessors.com and find that those who serve in more competitive and highly-ranked departments face more scrutiny in their comments than those who work in departments of lower status. The increased backlash for more prominently high-achieving women likely also comes in part from subverted gender roles, Rudman et al. [91] suggests. Over the course of five experiments, they pinpoint dominant gender roles as a key determinant of such backlash and find that backlash can be amplified when primed with a perceived threat of a high-ranking woman in a non-traditional sphere. Rudman et al. [91] also note the effect comes from both men and women. Taken together, these results explain some of our own data as likely being influenced by this combination of in-group protectionism and rigid mental gender roles. Even in our findings, this leads to women being penalized for attaining visible success in an academic setting.

While remaining largely congruent with existing literature, our study offers a few key additions. Our research contributes to the literature by providing experimental evidence exploring the relationship between gender and the use of professional titles and their impact on the perceived credibility of experts. Prior research has focused on one of these aspects, not their intersection. While the literature can establish a detectable contrast in the perceptions of female expert's credibility compared to those of men, it is not yet clear whether women suffer a higher credibility penalty from removing educational information signals such as professional titles. Furthermore, we present the first effort to explore a gender credibility gap through econometric analysis. Another limitation of the existing literature arises from the small nature of the surveys, with the aforementioned studies drawing conclusions from samples as small as 200 participants. Our work expands on the limited literature by investigating the gender credibility gap through the lens of perceived educational attainment. Additionally, we employ similar experimental designs and target audiences but with much larger sample sizes for more reliable results. Finally, we contribute to the literature by providing evidence of the understudied backlash effect women in higher status positions may face.

2.3 Experimental Design

Given that the effects of gender on perceived credibility are well documented as discussed above, we are instead interested in establishing whether there is a credibility effect of the use of professional titles that is different for men and women. Therefore, our survey utilized a 2x2 factorial design that randomly assigned article order, subject, and use of professional titles, while holding the expert's gender constant within survey versions. We constructed a survey instrument that presented respondents with news articles, each followed by several questions about their perceptions of the cited expert. In practice, survey respondents were presented with two article readings where the gender of the cited experts was the same for both readings, but the use of professional titles or honorifics varied. All respondents were presented with the same two articles. We sourced the articles from one major newspaper, one in the subject of economics and another in plant sciences, that cited an expert holding a doctoral degree and were published within the same week. We edited the articles into passages of about 350 words each. To ensure the credentials of the cited experts did not confound our experimental results, we provided the same information about each expert, referencing them as a "professor at an accredited U.S. university." Additionally, to vary the gender of the cited expert, we alternated between first names David and Susan and surnames Miller and Smith. David and Susan were each the third most popular baby names in 1964, while the Census 2000 records 73 percent of Smiths and 86 percent of Millers as White. We selected popular names that made the gender of the expert clear, while abstracting from potential confounding effects of varying their perceived race. This yielded four versions of each of the two articles, male/female with no professional title, and male/female with the professional title. Ultimately, each respondent was presented with a survey version with the two article readings where we randomized which article they read first, to minimize learning effects, and the use of the professional title between “Mr.”/ “Ms.” and “Dr.” while holding the gender constant. For example, if the first article presented to the participant referred to the expert as “Dr.,” then the second article referred to the expert as “Mr.” or “Ms.,” with first and surnames not repeated. We also included placebo tests, where neither gender nor use of professional titles varied. In this case, a respondent was presented with Dr. David Smith and Dr. Michael Miller, for example. We use the honorific “Ms.” as opposed to ”Mrs.” due to that preference exhibited by the AP Style Guidebook and the major newspaper used to source the articles. Although it is possible that respondents perceive an age effect to “Mrs.” compared to “Ms.”, given that this is the standard practice in news media, we believe this to be of low probability. We present an example of one of the articles in Figure 2.1.

The survey content was divided into three main sections and framed as an evaluation of “news article comprehension” to participants. to promote survey participation and completion, completion of the survey guaranteed at least one entry into a raffle to win one $15 Amazon gift card [92–94]. In the first section, as described above, participants were asked to read the two articles, each with nine accompanying questions evaluating different aspects of their perceived credibility of the cited expert on a seven-point Likert scale. Each news article subsection was preceded by a question evaluating the participant’s familiarity with the article topic to determine whether their familiarity with the subject had any effect on their perceptions of the cited experts. Participants were also asked a multiple-choice question regarding specifics of the article content prior to being shown the nine credibility-measuring questions as a quality check, ensuring that they had read the material. Almost all participants answered these questions correctly, at 95.54 percent. Participants correctly recalled the gender of the cited expert from the news articles 58.38 percent of the time. This compares to 60 percent in a similar study on source credibility by gender of authors [81]. Additionally, we can be confident that this is statistically different from participants simply guessing the gender of the experts as the survey question allowed for answers of “Don’t Know”, “Other”, and “Rather not say”, of which 23.10 percent selected, resulting in only 19 percent of respondents incorrectly recalling the experts’ gender. A simple t-test yields a p-value of 0.000 for the proportion of respondents correctly recalling the gender being statistically significantly different from 50 percent or guessing.

The second section of the survey consisted of three incentivized questions, a gender attitudes questionnaire, and a free response question asking participants what they considered was the purpose of the experiment. The incentivized questions were designed to elicit respondent’s perceptions on how other participants rated the credibility of the cited experts and to determine whether the treatment was salient. Due to our interest in truthful responses to these questions, each correct answer earned the participant an extra entry into a raffle to a gift card [95]. The gender attitudes questionnaire adapted from Swim [96] asked participants to rank their level of agreement with eight statements on a seven-point Likert scale. Finally, in the third and final section of the survey, we collected information on participants’ personal backgrounds. We collected their demographic information, including political identification between left- through right-leaning and education levels of their parents, as well information on news consumption habits. Again, we use this information to determine whether there exists any relationship between these demographic measures and their perceptions of the cited experts.

2.4 Data

The survey was administered to degree-seeking undergraduate students above the age of 18 at a large public university in the Washington, DC metro area. We recruited student participation through advertisements to their student emails. The university's registrar's office provided email access to all 20,444 degree-seeking undergraduates[1]. Students received the initial email advertisement on May 8, 2023, for an online economics experiment related to how students understood news, followed by two reminder emails one week apart. All students had three weeks to complete the survey from the receipt of the initial email. The survey closed on May 26, 2023. Our sample for analysis includes respondents that completed the survey. We also included quality control questions for each article by asking students to correctly recall some facts from each of the readings. Almost all participants, 96 percent, accurately recalled the information from the reading.

Table 2.1 shows a balance test of the demographic characteristics of respondents across title treatments and provides summary information of the participants in our sample. Unlike a standard balance table in Randomized Control Trials (RCTs) that compares baseline covariates across treatment and control groups, we present these demographic characteristics across the title and gender treatments that determine "treatment". Additionally, given that participants each read two articles, respondents may appear twice in these summary measures. For example, a student who was randomly assigned an article with a female Dr. followed by one with a male Dr. will appear in both of those columns. The average age across all respondents is 21.15 years old. This

[1]The first round of survey implementation began in November, 2022. However, in April 2023 we became aware that the registrar's office, in charge of disseminating our email advertisements to students, had inadvertently included information about the purpose of the study as gender-focused in the email signature notes. As a result, our data was potentially contaminated as at least some portion of respondents may have read the notes at the bottom of the email and responded to the survey with a particular gender-focused framing in mind. We reran the experiment the following month and acquired responses only from those who had not previously participated.

measure is overestimated as we only collect year of birth and the survey closed in May of 2023, resulting in our age calculations adding a year to age for those with birthdays after May. More than half of participants are female, at 57.49 percent. Only a minority of students indicated majors in economics and life sciences, the subjects of the survey articles. Additionally, 60.15 percent of students identify as leaning towards the political left. Finally, a majority of students, 84.01 percent, report their mothers attaining at least some college education and 82.87 percent report mothers working outside the home. Although four out of our nineteen demographic measures vary significantly across title treatments at the 10 percent significance level, results from testing for their joint orthogonality in Table 2.2 show that jointly, the covariates do not predict treatment. We perform this omnibus test to determine whether these demographic measures have any joint predictive power over respondents being treated and find the p-value of an F-test to be large at 0.66. Out of 549 students who opened the online survey, 394 students completed it. For most of our results which rely on a between-subjects empirical design, this yields a sample size of 788 observations. We address the potential bias from survey attrition in the Results section.

To measure how the use of professional titles may impact public perceptions of experts, we focus on their perceived credibility. [97] McCroskey and Young define credibility as relating to the attitude held by the recipient of information towards the source of the information. Although some researchers, such as Armstrong and McAdams [81], have used scales from 1-3 to rank trustworthiness, a similar attribute to credibility, we use a 7-point Likert scale for consistency with most literature and to increase measurement precision by increasing our ability to measure degrees of opinion. Paul, Sui, and Searles [80], Nagle, Brodsky and Weeter [98], Zhu, Aquino, and Vadera [99], and Boczek, Dogruel and Schallhorn [100], among others, also use some form of a Likert scale to measure different facets, such as likability, confidence, and knowledge of the intersection between gender and perceived credibility. Indeed, a benefit of the usage of the Likert scale is how it allows for separate measurement of such aspects on different component sub-scales that can be formed into a composite measure used to quantify credibility [83]. Therefore, our dependent variable of interest, perceived credibility of the referenced expert by readers, takes the form of an index of responses to five related questions, which we denote as the credibility score. It is generated using responses from questions asking participants their level of agreement with statements about whether the cited expert is a respected scholar in their field, works at one of the best universities in the U.S., expresses his- or herself with clarity, provides reliable information, and provides trustworthy information. The responses are in the form of a Likert scale, so for ease of interpretation, we construct the composite measure as an equally weighted average of z-scores of each of the five components, calculated by subtracting the mean and dividing by the standard deviation of the control group, and also standardizing the resulting index [101]. Therefore, each component is mean zero with standard deviation one for the control group, and the index has similar properties. For both the components and the index, higher values mean more credible perceptions of the cited expert. As a measure of the internal consistency of the use of these five prompts as a measure of credibility, we calculate the inter-item covariances and estimate Cronbach's α = 0.83. This confirms our credibility index is scale reliable and the credibility prompts are highly related as they meet the general rule of thumb of 0.70.

2.5 Empirical Framework

Due to the randomization of treatment assignment in our experiment, treatment is at the article level such that treated articles are those using the professional title "Dr." and our control group consists of articles using honorifics such as "Mr." or "Ms." to refer to the cited experts. To estimate the effect of expert gender, we thus compare article responses that were assigned female experts compared to those assigned to male experts. We begin our analysis by estimating the relationship between professional titles, gender, and expert credibility with a within-subjects design. This approach allows us to abstract from any unobservable characteristics across individuals that may systematically differ across treatment assignment and could confound our estimate of the causal impact of professional titles and gender on expert credibility. We estimate the effect of the use of the professional title "Dr." compared to the honorific "Mr." or "Ms." across genders through the following model:

$$Credibility_{is} = \alpha + \beta Dr_{is} + \gamma \boldsymbol{Dr}_{is} \times \textbf{\textit{Female\_Expert}}_i + \delta_i + \Gamma X_i + \epsilon_{is} \tag{2.1}$$

Where for each respondent i and article s, $Credibility_{is}$ is the composite credibility index, $\boldsymbol{Dr}_{is}$ indicates treatment as the use the professional title in the article text and $Female\_Expert_i$ indicates whether the expert in the news article is female. Recall that the gender of the cited experts does not differ by article, but by individual, where no respondent is presented with two articles that vary gender. Thus, our results from this model cannot estimate the main effect of the expert's gender on perceived credibility. We include individual fixed effects δ_i to capture within-subject variation. Finally, X_{is} is a vector of article controls including the subject of the article and the reading order of the article.

Next, we estimate the differential impact of professional titles across gender treatments using a between-subjects design. Now our treatment group consists of participant-article observations with articles using professional titles, and our control group is comprised of participant-level observations with articles using only honorifics. The model is:

$$Credibility_{is} = \alpha + \beta_1 \boldsymbol{Dr}_{is} + \beta_2 \boldsymbol{Expert\_Gender}_{is} + \beta_3 \boldsymbol{Dr}_{is} \times \boldsymbol{Expert\_Gender}_{is}$$
$$+\Gamma X_{is} + \Phi W_i + \epsilon_{is} \qquad (2.2)$$

Where all parameters are measured as before, but we now include W_i as a set of respondent characteristics. These characteristics were collected at the end of the survey and include age, gender, ethnicity, prior research experience, major, parents' education and whether their mom worked outside of the home, political standing, class standing, news reading frequency, and graduate education plans. Our main parameter of interest β_3 estimates the differential impact of professional titles for female experts compared to male experts. The coefficients β_1 and β_2 measure the effect of the information signal of a professional title on the reader's perceived credibility of the quoted expert and the effect of the expert's gender, respectively. A $\beta_1 > 0$ indicates a positive title effect on perceived credibility while a $\beta_2 < 0$ indicates a gender penalty.

We present results from a few other specifications to test for potential confounders. First, we restrict the sample to only responses from the first article respondents read, as learning and fatigue effects could affect their responses to the same questions in the second article. Next, we re-estimate Equation 2.4 separately by participant gender, male and female, to explore heterogenous treatment effects. We also present results for each of the five credibility score components to ascertain whether any particular component drives our results. Finally, we asked participants to read one of the selected quotes from the cited experts in each article and determine whether they believed the quote to be based on expertise or opinion and present those results below.

2.6 Results

Results from the within-subjects model in Equation 2.1 are presented in Table 2.3. Column (2) shows our estimates of the effects of professional titles across genders including article-level covariates. We establish a statistically significantly positive title effect on perceived credibility of cited experts. The estimated parameter shows an increase of 0.138 standard deviations attributable to the use of professional titles. However, we are unable to detect a differential impact across gender with the small and statistically insignificant estimate of $\gamma = 0.049$. The inclusion of the article-level controls across columns (1) and (2) only slightly increases the coefficient on professional titles for the interaction term, suggesting that most of the variation is captured by our individual fixed effects. Additionally, neither of the included article controls, whether it was an economics or plant sciences article, or whether it was the first article the respondent read, have any impact on the perceived credibility of the cited experts.

Table 2.4 presents the estimation results of our between-subjects model in Equation 2.2, with column (3) showing results including article and individual level controls. The unit of analysis for these and the remaining results is the article-response. Consequently, respondents appear twice in the sample, once for each of their article responses. The results in Table 2.4 show that while not statistically significant, we still observe a positive title effect for expert credibility of 0.050 standard deviations. However, we remain unable to detect a differential effect for female experts. Nonetheless, the signs of the point estimates suggest that while women are perceived as less credible, the use of professional titles offsets this by almost half. The estimated effect of a female expert is -0.078 standard deviations while that for female experts with a professional title is 0.058 standard deviations, leaving only a 0.020 standard deviation gap. Additionally, a test of whether the sum of the coefficients on the interaction term and on whether the expert is female is equal to zero yields a p-value of 0.81. Thus, this suggests that the use of titles for women offsets the gender credibility they face. For the included individual-level covariates, only whether the student is a life sciences major has a statistically significant impact on their perceived credibility of the cited experts. Being a life sciences major results in lower credibility ratings by 0.174 standard deviations.

Next, we present the re-estimation of Equation 2.2 restricting the sample to responses to the first article read by all participants in Table 2.5 and we investigate possible heterogenous treatment effects across respondents' gender. We enact the sample restriction of only responses to the first article read by respondents to investigate whether learning effects or fatigue as participants move to the second article play any role in our results. The sample restriction yields noisy estimates of our coefficient of interest as the sample size is halved to 391 respondents and our standard errors substantially increase. While we still observe a negative effect on credibility for female experts and a positive effect for women referenced with the title "Dr.", we now observe a noisy negative estimate for the title effect on credibility. However, we again are unable to reject the null hypothesis that the sum of the coefficient on the interaction term and on whether the cited expert is female is equal to zero across all specifications, suggesting the use of titles offset the gender credibility penalty. For example, the resulting p-value for the coefficients in column (3) is 0.96. Given that we observe no substantial changes in the estimated coefficients, we are not concerned that learning or fatigue effects confounded our main estimates. These results confirm statistically insignificant, close to zero estimates for the differential impact of professional titles on expert credibility across gender. Table 2.6 estimate Equation 2.2 separately for female and male respondents in Columns (1) and (2), respectively. While the general patterns persist of no detectable differential credibility effects across genders, it is interesting to note that female respondents rate female experts as less credible than male respondents. These results are consistent with the own-rating and same gender literature previously discussed where along several components, women rate themselves and other women more poorly than men rate themselves. Nevertheless, for female respondents, we still observe that the use of professional titles for women mostly recovers the gender penalty. On the other hand, for male respondents, while they do not rate women as less credible, they do in fact impose a penalty to women referenced with a professional title. This is consistent with the backlash effect discussed in the literature previously referenced where women in positions of power or prestige can suffer from increased scrutiny from their peers and subordinates.

We also present the results of Equation 2.2 by credibility index component in Table 2.7 to investigate which components may be driving our results. The table presents results for each of the five credibility questions individually. Again, estimates suggest no differential credibility penalty for female experts compared to male experts. All but one measure yield positive point estimates for the effect of the use of professional titles, and all but the same measure yield negative point estimates for the effect of being female. This pattern persists for the coefficients on the interaction term. Overall, we consistently observe small positive gains in credibility from the use of titles that compensate for the gender penalty for women. Figures 2.2 and 2.3 graphically present the distribution of responses to each of the five core credibility questions by gender and title treatments. Each figure also reports the corresponding p-value for the Kolmogorov-Smirnov test of equality of distributions. The figures show that respondents are more likely to agree to the cited experts being credible. One exception is the question about whether the cited expert is at one of the best universities, where most respondents report being unsure. These figures also show that the mass of the distribution of responses does not belong to the extreme endpoints, assuaging concerns on survey manipulation by respondents. Finally, we also present results from asking participants whether they considered a quote in the text from the experts as based on expertise or opinion. While the results in Table 2.8 do not show significant title or gender effects, we do observe a marginally statistically significant effect of titles for women, where respondents find that female experts addressed using “Dr.” are more likely to be considered as making statements based on their expertise as opposed to their opinions. The estimated coefficient represents an effect of 0.119 standard deviations.

In our online experiment, we experienced a high rate of attrition at 28.23 percent. Unlike many RCT designs, we were unable to collect baseline demographic information about survey participants that we could subsequently use to examine rates of attrition. In our survey design, the demographic module occurred at the end of the online survey. This rules out using techniques such as Inverse Probability Weighting (IPW) or selective attrition tests like Ghanem, Hirshleifer and Ortiz-Becerra [102]. However, we are able to determine whether respondents’ random treatment assignment is correlated with their probability of attrition. We present these calculations in Table 2.9. Attrition rates across the title-gender treatments are almost identical except for the Male “Dr.” treatment assignment that is slightly lower at 25.48 percent. Nevertheless, we observe only a minuscule correlation with attrition, at -0.027. Consequently, attrition was orthogonal to treatment assignment.

Nevertheless, we proceed by estimating lower and upper bounds for our main between-subjects specification from Equation 2.2 using Horowitz-Manski bounds [103] in Table 2.10. Interestingly, neither the bounds for the gender effect on credibility nor the differential effect of titles on gender include zero within the bounds. Nonetheless, for either of these estimates, the upper bound remains statistically insignificantly different from zero. However, while the bounds for the title effect are too large to provide any additional insights, the bounds around both the gender and interaction terms are narrow, within 0.1 standard deviations. These results suggest that at most, we could expect a gender penalty of 0.103 standard deviations and a positive title effect for women of 0.071 standard deviations. Another bounding exercise, proposed by David Lee [104], and commonly referred to as Lee bounds, would not be informative in this context. Lee bounds require the additional assumption of monotonicity, whereby treatment can only affect attrition in one direction. As shown in Table 2.9, treatment assignment of the use of a professional title is positively correlated with attrition for female experts but negative for male experts. Therefore, we do not implement this procedure.

2.7 Discussion

We investigate the possibility of a gender credibility penalty for female experts when not referenced by their earned professional titles in major news outlets. We hypothesize that given documented gender bias against women in the media, workplace environments, and the media, among other realms, the removal of information about their training, namely the use of a professional title, as is common practice in many major news outlets, could exacerbate and even aggravate existing biases. We design and implement an online survey experiment eliciting perceptions from undergraduate students of the credibility of experts cited in major news articles when the title ”Dr.” or an honorific such as ”Ms.” is used. The results from this experiment suggest that while there exists a positive effect of the use of professional titles on expert’s credibility, we are unable to detect either a significant gender penalty for women or a differential title effect across genders. Our estimated effects are small and close to zero. This may result from a number of reasons, such as sample size restrictions, high variance in survey responses, treatment salience, or a true null effect. Our ex-ante power calculations referencing a somewhat similar study [105] indicated we could detect a treatment effect of 0.21 standard deviations with a sample of 600 respondents. So, while we were powered to detect a moderately-sized treatment effect, our estimates are much smaller. Ex-post power calculations indicate a sample size of over 6,000 students are needed to detect an effect of our size. Perhaps the main concern is that treatment was not salient. With only 58 percent of respondents able to correctly recall the gender of the cited expert at the end of the experiment, this raises doubts whether respondents paid attention to the gender, beyond the title used to reference them. If this is the case, it could be that the true causal effect of the use of professional titles does in fact result in a differential impact across genders. As potential future research, we are interested in implementing the online survey experiment again in multiple waves across different sample populations to increase sample size and be able to better study heterogeneous treatment effects. One modification we could implement would be to edit the article texts to make the gender of the experts more salient, by using pronouns more often and including the expert's names in the prompts, for example. Ultimately, there may be important dimensions along which perceptions of female experts compared to their male experts vary in meaningful ways beyond their perceived credibility that we could pursue estimating with our survey design.

Figure 2.1: Example article

UNIVERSITY OF MARYLAND

For Disabled Workers, a Tight Labor Market Opens New Doors

The strong late-pandemic labor market is giving a lift to a group often left on the margins of the economy: workers with disabilities.

Employers, desperate for workers, are reconsidering job requirements, overhauling hiring processes and working with nonprofit groups to recruit candidates they might once have overlooked. At the same time, companies' newfound openness to remote work has led to opportunities for people whose disabilities make in-person work difficult or impossible.

As a result, the share of disabled adults who are working has soared in the past two years, far surpassing its pre-pandemic level and outpacing gains among people without disabilities, says Dr. David Smith, a professor at an accredited U.S. university. More than 35 percent of disabled Americans ages 18 to 64 had jobs in September 2022, up 4 percentage points just before the pandemic.

People with disabilities report that they are getting not only more job offers, but better ones, with higher pay, more flexibility and more openness to providing accommodations that once would have required a fight, if they were offered at all.

“The new world we live in has opened the door a little bit more,” said Dr. Smith. “The doors are opening wider because there's just more demand for labor.”

Workers with disabilities — like other groups that face obstacles to employment, such as those with criminal records — tend to benefit disproportionately from strong job markets, when employers have more of an incentive to seek out untapped pools of talent. But when recessions hit, those opportunities quickly dry up.

“We have a last-in, first-out labor market, and disabled people are often among the last in and the first out,” said Dr. Smith.

Remote work, however, has the potential to break that cycle, at least for some workers. Dr. Smith says that employment has risen for workers with disabilities across industries as the labor market improved, consistent with the usual pattern. But it has improved especially rapidly in industries and occupations where remote work is more common. And this shift toward remote work, unlike the red-hot labor market, is likely to prove lasting.

“Disabled adults have seen employment rates recover much faster,” Dr. Smith said. “Not only is this a permanent thing, but it's going to improve.”

Employers that don't find ways to accommodate workers with disabilities — whether through remote work or other adjustments — are going to continue to struggle to find employees, said Dr. Smith, “employers have to pivot. Otherwise, this labor shortage may be more permanent.”

Figure 2.2: Responses by Title Treatment

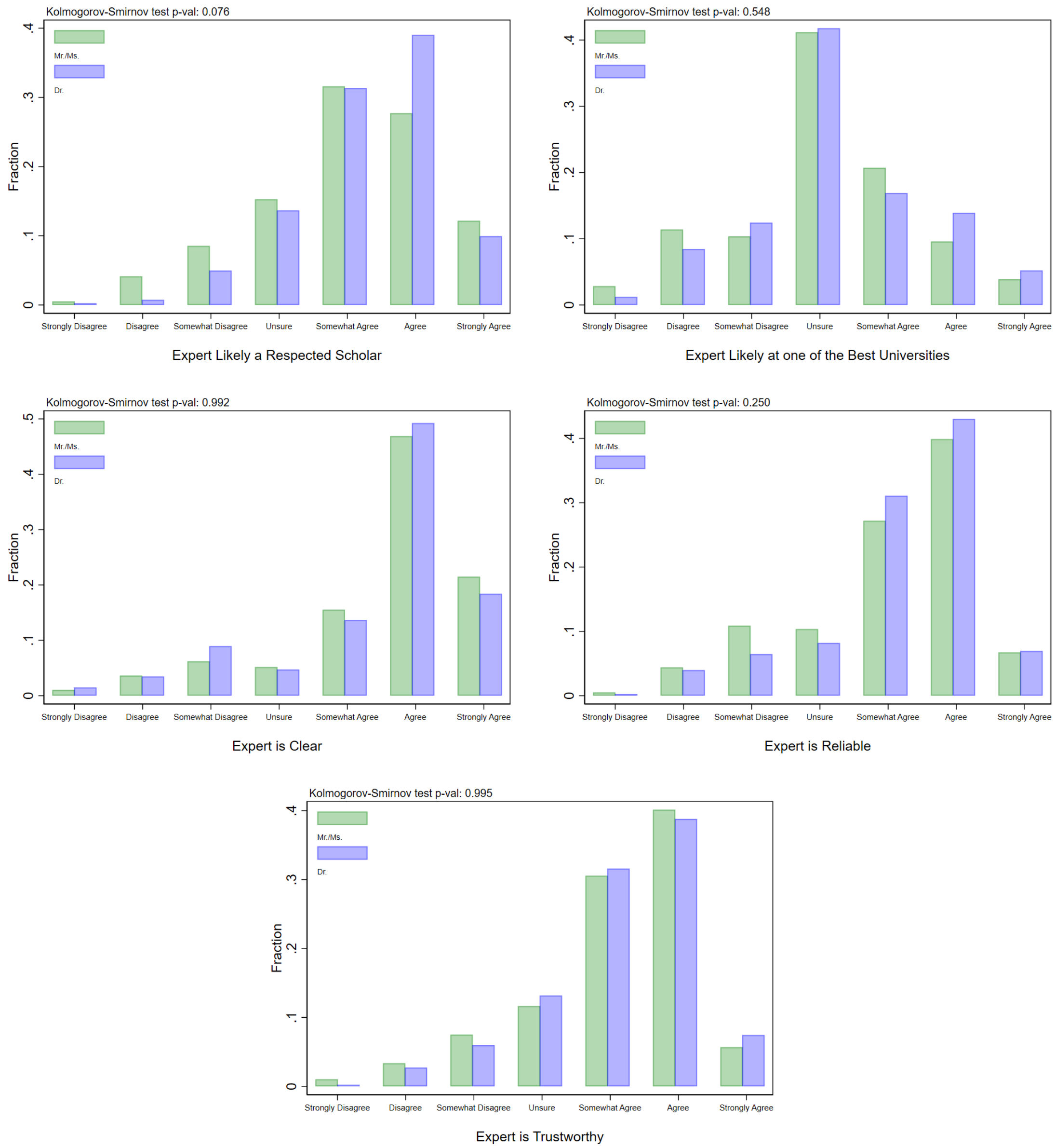


Figure 2.3: Responses by Gender Treatment

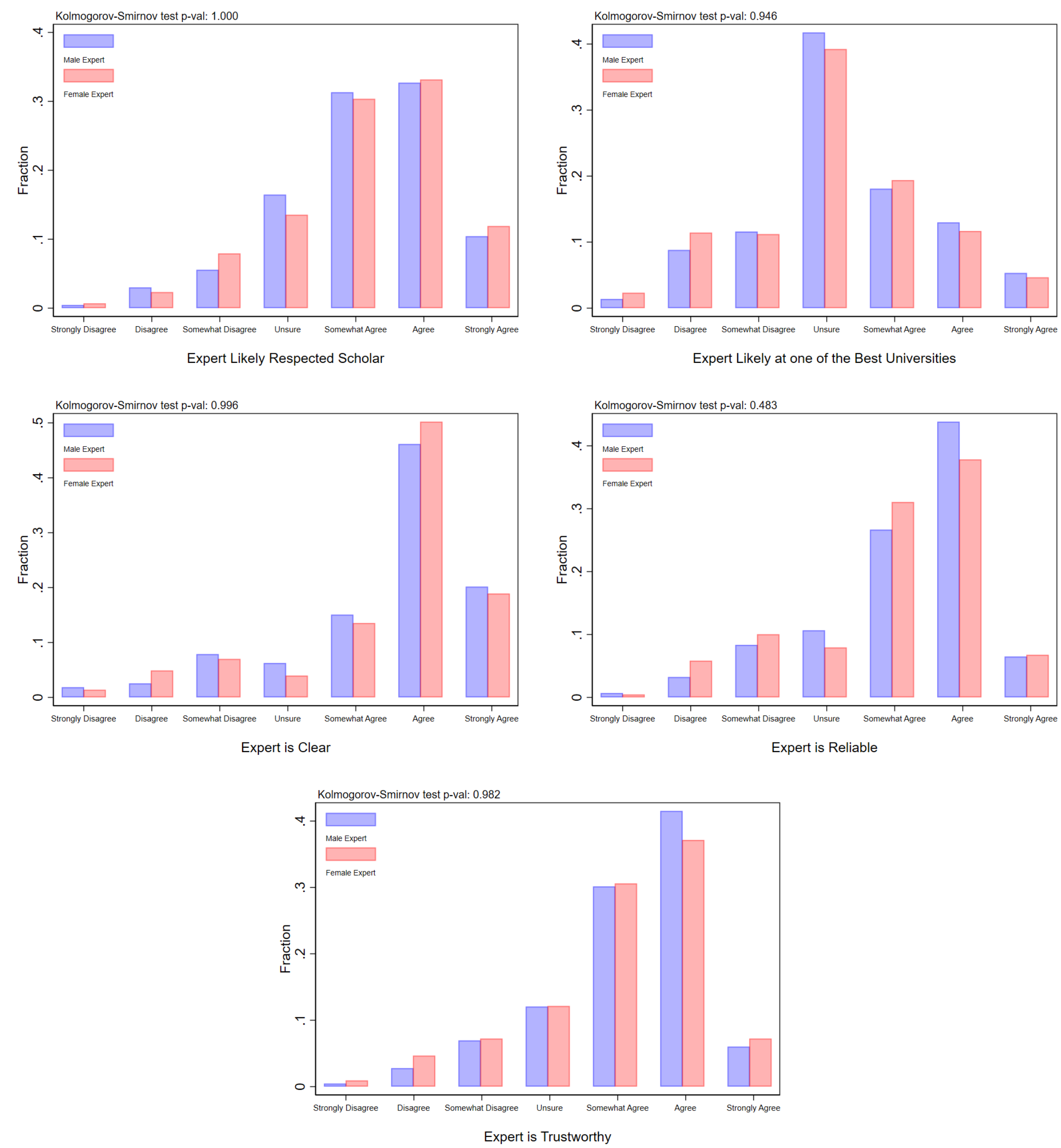


Table 2.1: Summary Statistics

	Article Version				
	Mr.	Ms.	Female Dr.	Male Dr.	P-value
Demographic Characteristics					
Age	21.14	21.23	21.30	21.00	0.61
	(2.39)	(2.78)	(2.69)	(2.41)	
Female	0.54	0.60	0.58	0.53	0.70
	(0.50)	(0.49)	(0.49)	(0.50)	
Black	0.07	0.14	0.15	0.06	0.06
	(0.25)	(0.35)	(0.36)	(0.24)	
Hispanic/Latino	0.09	0.12	0.12	0.12	0.53
	(0.29)	(0.33)	(0.32)	(0.32)	
Prior Experiment Experience	0.75	0.73	0.78	0.80	0.07
	(0.43)	(0.45)	(0.42)	(0.40)	
Weekly News Reader +	0.46	0.41	0.41	0.47	0.70
	(0.50)	(0.49)	(0.49)	(0.50)	
University Major					
Economics or business school major	0.19	0.18	0.19	0.19	0.99
	(0.39)	(0.39)	(0.40)	(0.39)	
Life sciences major	0.16	0.21	0.21	0.14	0.47
	(0.37)	(0.41)	(0.41)	(0.35)	
Parents' education and work					
Mother - Some College +	0.77	0.87	0.87	0.80	0.16
	(0.42)	(0.34)	(0.34)	(0.40)	
Father - Some College +	0.78	0.88	0.84	0.80	0.06
	(0.42)	(0.32)	(0.37)	(0.40)	
Mother - Worked Outside Home	0.83	0.81	0.80	0.84	0.87
	(0.38)	(0.39)	(0.40)	(0.37)	
Political Identification					
Left-Leaning	0.57	0.58	0.59	0.58	0.97
	(0.50)	(0.50)	(0.49)	(0.49)	
Right-Leaning	0.09	0.09	0.08	0.07	0.69
	(0.29)	(0.28)	(0.28)	(0.26)	
Class Standing					
First year	0.26	0.27	0.22	0.29	0.08
	(0.44)	(0.45)	(0.42)	(0.46)	
Second year	0.25	0.24	0.29	0.28	0.42
	(0.43)	(0.43)	(0.45)	(0.45)	
Third year	0.28	0.27	0.30	0.23	0.27
	(0.45)	(0.44)	(0.46)	(0.42)	
Fourth year	0.19	0.20	0.18	0.18	0.83
	(0.40)	(0.40)	(0.38)	(0.38)	
Fifth year +	0.02	0.02	0.01	0.02	0.76
	(0.14)	(0.14)	(0.12)	(0.14)	
Observations	145	147	146	154	

Notes: Standard errors are in parentheses. Observations are at the individual-article level and include our core analysis group consisting of respondents who completed the survey. The last column presents a p-value from a test of the joint equality of the means across article treatments.

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table 2.2: Omnibus Test

	Article Version			
	Mr.	Ms.	Female Dr.	Male Dr.
Age	0.019	0.001	-0.003	-0.006
	(0.019)	(0.018)	(0.017)	(0.022)
Female	-0.089	-0.147	0.055	-0.022
	(0.081)	(0.089)	(0.083)	(0.084)
Black	-0.152	0.076	-0.099	0.178
	(0.165)	(0.131)	(0.108)	(0.206)
Hispanic/Latino	0.076	0.197	-0.263**	0.131
	(0.135)	(0.139)	(0.126)	(0.128)
Prior Experiment Experience	0.144	-0.011	0.244**	0.001
	(0.092)	(0.098)	(0.094)	(0.096)
Economics or business school major	0.024	-0.154	0.148	0.005
	(0.111)	(0.108)	(0.106)	(0.110)
Life sciences major	-0.292**	-0.046	0.111	0.198
	(0.118)	(0.111)	(0.103)	(0.126)
Mother - Some College +	-0.062	-0.102	0.214	0.145
	(0.116)	(0.145)	(0.148)	(0.121)
Father - Some College +	0.169	0.103	-0.419***	-0.048
	(0.114)	(0.156)	(0.137)	(0.121)
Mother - Worked Outside Home	0.134	-0.110	0.035	0.031
	(0.112)	(0.103)	(0.099)	(0.096)
Left-Leaning	-0.126	-0.020	0.025	0.098
	(0.093)	(0.097)	(0.088)	(0.097)
Right-Leaning	-0.117	0.109	-0.119	-0.175
	(0.150)	(0.160)	(0.150)	(0.137)
First year	0.508*	0.174	0.044	-0.037
	(0.267)	(0.364)	(0.381)	(0.304)
Second year	0.252	0.139	0.331	0.112
	(0.267)	(0.360)	(0.372)	(0.296)
Third year	0.297	0.125	0.367	-0.126
	(0.265)	(0.349)	(0.365)	(0.295)
Fourth year	0.319	-0.021	0.340	-0.053
	(0.264)	(0.357)	(0.366)	(0.294)
Weekly News Reader +	-0.074	-0.037	-0.000	0.090
	(0.083)	(0.087)	(0.081)	(0.082)
P-value	0.091	0.369	0.621	0.212
N	145	145	146	151

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Observations are individual-article level. Standard errors are clustered at the individual level and shown in parentheses. P-value is from a test of joint significance of the coefficients.

Table 2.3: "Dr." Effect: Within-Subjects Model

	Credibility Score	
	(1)	(2)
Dr	0.137*	0.138*
	(0.072)	(0.072)
Dr x Female Expert	0.048	0.049
	(0.110)	(0.110)
Economics Article		0.030
		(0.036)
First Article Read		-0.029
		(0.036)
N	780	780
Adj. R^2	0.56	0.56

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Standard errors are clustered by survey respondent and are in parentheses. Observations are at the individual-article level. The outcome variable is a mean zero credibility score index constructed as the average of the z-scores of five credibility-related survey questions such that higher scores mean higher credibility ratings.

Table 2.4: "Dr." Effect: Between-Subjects Model

	Credibility Score		
	(1)	(2)	(3)
Dr	0.074	0.075	0.050
	(0.075)	(0.075)	(0.078)
Female Expert	-0.056	-0.056	-0.078
	(0.089)	(0.089)	(0.090)
Dr x Female Expert	0.035	0.035	0.058
	(0.110)	(0.110)	(0.111)
Economics Article		0.030	0.029
		(0.036)	(0.036)
First Article Read		-0.030	-0.030
		(0.036)	(0.036)
Female Respondent			0.056
			(0.069)
Econ or Business School Major			0.102
			(0.080)
Life Sciences Major			-0.174*
			(0.097)
Mother - Some College +			0.035
			(0.111)
Father - Some College +			-0.055
			(0.103)
Mother - Worked Outside Home			-0.061
			(0.098)
Weekly News Reader +			0.023
			(0.071)
N	788	788	788
Individual Controls			✓
Adj. R^2	0.00	-0.00	0.01

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Standard errors are clustered by survey respondent and are in parentheses. Observations are at the individual-article level. The outcome variable is a mean zero credibility score index constructed as the average of the z-scores of five credibility-related survey questions such that higher scores mean higher credibility ratings. In addition to the individual level controls presented above, we include respondents' age, race, political leaning, class standing, and prior experimental research experience.

Table 2.5: “Dr.” Effect: Between-Subjects Model: First Article

	Credibility Score		
	(1)	(2)	(3)
Dr	0.010	0.005	-0.068
	(0.099)	(0.098)	(0.104)
Female Expert	-0.095	-0.100	-0.131
	(0.113)	(0.112)	(0.115)
Dr x Female Expert	0.070	0.077	0.136
	(0.150)	(0.150)	(0.154)
Economics Article		0.092	0.101
		(0.074)	(0.073)
Female Respondent			0.084
			(0.076)
Econ or Business School Major			0.130
			(0.091)
Life Sciences Major			-0.221**
			(0.112)
Mother - Some College +			0.021
			(0.121)
Father - Some College +			-0.159
			(0.117)
Mother - Worked Outside Home			-0.024
			(0.118)
Weekly News Reader +			0.019
			(0.080)
N	391	391	391
Individual Controls			✓
Adj. R^2	-0.005	-0.003	0.006

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Standard errors are clustered by survey respondent and are in parentheses. The outcome variable is a mean zero credibility score index constructed as the average of the z-scores of five credibility-related survey questions such that higher scores mean higher credibility ratings. Observations are at the individual-article level. In addition to the individual level controls presented above, they include respondents’ age, race, political leaning, class standing, and prior experimental research experience. Sample includes responses to solely the first article read by each participant.

Table 2.6: “Dr.” Effect: Between-Subjects Model: By Respondent Gender

	Credibility Score	
	Female Respondents	Male Respondents
	(1)	(2)
Dr	0.048	0.142
	(0.111)	(0.102)
Female Expert	-0.163	0.059
	(0.124)	(0.130)
Dr x Female Expert	0.103	-0.075
	(0.156)	(0.159)
Economics Article	0.040	0.031
	(0.049)	(0.057)
First Article Read	0.010	-0.059
	(0.050)	(0.057)
Econ or Business School Major	0.144	0.073
	(0.107)	(0.141)
Life Sciences Major	-0.212*	0.011
	(0.122)	(0.151)
Mother - Some College +	-0.047	0.106
	(0.142)	(0.191)
Father - Some College +	0.091	-0.087
	(0.138)	(0.199)
Mother - Worked Outside Home	0.024	-0.146
	(0.127)	(0.119)
Weekly News Reader +	-0.102	0.145
	(0.107)	(0.094)
N	453	308
Individual Controls	✓	✓
Adj. R^2	0.030	0.024

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Standard errors are clustered by survey respondent and are in parentheses. The outcome variable is a mean zero credibility score index constructed as the average of the z-scores of five credibility-related survey questions such that higher scores mean higher credibility ratings. Observations are at the individual-article level. In addition to the individual level controls presented above, we include respondents’ age, race, political leaning, class standing, prior experimental research experience, graduate study plans, and news reading frequency.

Table 2.7: "Dr." Effect: Between-Subjects Model: By Credibility Score Component

	Respected Scholar	Best University	Clarity	Reliability	Trustworthiness
	(1)	(2)	(3)	(4)	(5)
Dr	0.136	0.091	-0.041	0.051	0.012
	(0.097)	(0.096)	(0.112)	(0.093)	(0.094)
Female Expert	-0.031	-0.086	0.019	-0.170	-0.121
	(0.113)	(0.117)	(0.115)	(0.107)	(0.110)
Dr x Female Expert	0.137	0.056	-0.075	0.116	0.055
	(0.136)	(0.141)	(0.156)	(0.137)	(0.136)
Economics Article	0.031	0.071*	0.086	-0.028	-0.013
	(0.053)	(0.043)	(0.053)	(0.056)	(0.050)
First Article Read	0.101*	-0.055	-0.062	-0.112**	-0.024
	(0.053)	(0.043)	(0.053)	(0.056)	(0.050)
Female Respondent	0.165**	0.151*	0.023	-0.027	-0.032
	(0.076)	(0.091)	(0.102)	(0.082)	(0.085)
Econ or Business School Major	0.084	0.074	0.175	0.151	0.026
	(0.088)	(0.116)	(0.115)	(0.100)	(0.107)
Life Sciences Major	-0.219*	-0.321**	0.135	-0.249**	-0.216*
	(0.112)	(0.124)	(0.122)	(0.112)	(0.122)
Mother - Some College +	-0.102	0.104	0.061	0.105	0.006
	(0.130)	(0.151)	(0.126)	(0.127)	(0.139)
Mother - Worked Outside Home	-0.034	-0.061	-0.023	-0.098	-0.089
	(0.111)	(0.115)	(0.124)	(0.111)	(0.117)
Weekly News Reader +	-0.072	-0.018	0.063	0.040	0.102
	(0.081)	(0.098)	(0.094)	(0.084)	(0.088)
N	788	788	788	788	788
Article Controls	✓	✓	✓	✓	✓
Individual Controls	0.022	0.036	-0.006	0.026	0.013

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Standard errors are clustered by survey respondent and are in parentheses. Outcomes are a z-score of the reponses to each of the survey prompts. Observations are at the individual-article level. In addition to the individual level controls presented above, we include respondents' age, race, political leaning, class standing, prior experimental research experience, graduate study plans, and news reading frequency.

Table 2.8: “Dr.” Effect: Between-Subjects Model - Expertise or Opinion?

	Credibility Score		
	(1)	(2)	(3)
Dr	0.018	0.021	0.010
	(0.051)	(0.048)	(0.047)
Female Expert	-0.055	-0.059	-0.051
	(0.050)	(0.049)	(0.048)
Dr x Female Expert	0.103	0.105	0.119*
	(0.073)	(0.068)	(0.067)
Economics Article		-0.364***	-0.362***
		(0.033)	(0.033)
First Article Read		-0.065**	-0.064*
		(0.033)	(0.033)
Female Respondent			-0.009
			(0.035)
Econ or Business School Major			-0.026
			(0.047)
Life Sciences Major			-0.023
			(0.044)
Mother - Some College +			-0.063
			(0.054)
Father - Some College +			-0.046
			(0.048)
Mother - Worked Outside Home			-0.118***
			(0.045)
Weekly News Reader +			0.012
			(0.036)
N	777	777	777
Individual Controls			✓
Adj. R^2	0.004	0.138	0.152

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Standard errors are clustered by survey respondent and are in parentheses. Observations are at the individual-article level. Sample includes respondents that finished the survey. In addition to the individual level controls presented above, we include respondents’ age, race, political leaning, class standing, and prior experimental research experience. The outcome variable is a binary indicator for whether the respondent perceived the cited expert’s quote as based on expertise, instead of opinion.

Table 2.9: Attrition by Article Treatment Assignment

	Mr.	Ms.	Female Dr.	Male Dr.
	(1)	(2)	(3)	(4)
Attrited	0.024	0.024	0.038	-0.027
	(0.040)	(0.040)	(0.040)	(0.039)
Attrition Rates	28.64%	28.64%	29.47%	25.48%
Observations	549	549	549	549

Notes: Observations at the individual level. Regressions include control for which article participant was presented first. Standard errors are clustered at the respondent level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.10: "Dr." Effect: Between-Subjects Model: Horowitz-Manski Bounds

	Credibility Score		
	Lower Bound	Main Specification	Upper Bound
	(1)	(2)	(3)
Dr	-1.277***	0.050	1.344***
	(0.139)	(0.078)	(0.138)
Female Expert	-0.103	-0.078	-0.001
	(0.126)	(0.090)	(0.128)
Dr x Female Expert	0.024	0.058	0.071
	(0.199)	(0.111)	(0.194)
Economics Article	0.032	0.029	0.006
	(0.072)	(0.036)	(0.069)
First Article Read	0.047	-0.030	-0.105
	(0.072)	(0.036)	(0.069)
Econ or Business School Major	0.085	0.102	0.118
	(0.101)	(0.080)	(0.104)
Life Sciences Major	-0.224*	-0.174*	-0.113
	(0.118)	(0.097)	(0.121)
Mother - Some College +	0.050	0.035	0.027
	(0.141)	(0.111)	(0.131)
Father - Some College +	-0.166	-0.055	0.055
	(0.124)	(0.103)	(0.139)
Mother - Worked Outside Home	-0.073	-0.061	0.007
	(0.115)	(0.098)	(0.122)
Weekly News Reader +	0.029	0.023	0.028
	(0.088)	(0.071)	(0.087)
N	1,098	788	1,098
Individual Controls	✓	✓	✓
Adj. R^2	0.241	0.010	0.289

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Standard errors are clustered by survey respondent and are in parentheses. Observations are at the individual-article level. The outcome variable is a mean zero credibility score index constructed as the average of the z-scores of five credibility-related survey questions such that higher scores mean higher credibility ratings. In addition to the individual level controls presented above, we include respondents' age, race, political leaning, class standing, prior experimental research experience, graduate study plans, and news reading frequency. The upper and lower bound estimates are calculated implementing Horowitz-Manski bounds to address survey attrition.

Chapter 3: Increasing Supply of Physicians and Birth Outcomes

3.1 Introduction

Expanding health care access to vulnerable populations remains a challenging task for many countries. Brazil, a country that constitutionalized the right to healthcare in 1988, still exhibits large inequalities in access to healthcare across demographic and geographic characteristics. Compared to other OECD countries, Brazil has lower levels of physicians, but its physicians are also concentrated in metropolitan regions. In 2012, Brazil had on average 1.8 primary care physicians per 1,000 residents, with 22 out of 27 states falling below that level. For comparison, in the same year, the UK and Argentina had 2.7 and 3.9 physicians per 1,000 residents, respectively [106]. However, not all Brazilian regions suffered from physician shortages. The capital city Brasilia had 3.6 primary care physicians per 1,000 residents, while over 1,900 municipalities that had less than one primary care physician per 3,000 residents [107].[1] Consequently, although Brazil universalized healthcare decades ago, it has faced serious challenges in implementing access in a way that is equitable for all its citizens. Physicians generally preferred work in the capitals and larger cities, contributing to this significant difference in healthcare access between the more remote and rural areas to the north of the country and the urbanized Southern and Southeastern regions of the country.

[1] A Brazilian municipality can be thought of as similar to a US county.

To address the shortage of physicians in underserved areas, and more generally to increase average numbers of primary care physicians, Brazil introduced the Mais Medicos Program (PMM), translating to ”More Doctors”. Doctors under Mais Medicos work for the Sistema Único de Saúde (SUS), or the unified health system. SUS is the publicly-funded national health care system that was created in 1988. SUS is present in all 5,000+ Brazilian municipalities, and according to the Ministry of Health, comprises 80 percent of medical appointments in the country. PMM emphasizes three goals: emergency hiring of doctors to address the national shortage of healthcare workers, increasing slots in medical programs around the country, and updating medical training curricula to focus on “humanized” care and the importance of primary care. This analysis will focus on the first goal, that of increasing the supply of health workers throughout the country.

3.2 Context

The Mais Medicos program was announced by President Rousseff in June 2013 following mounting public pressure and protests over the massive federal spending on preparation for the World Cup and Summer Olympics soon to take place in Rio de Janeiro. Kirk, Kirk and Walker [108] argue that the program was created most likely as a political tactic to appease public concern on social disparities, given that President Rousseff was up for election the following year. The shortage of doctors in Brazil was significant, with at least 12 million people in municipalities with no doctors [109]. Prior to Mais Medicos, in 2011 the government issued a public notice seeking doctors for positions in the North and Northeast of the country, where they were the most scarce, through the Programa de Valorização da Atenção Básica (PROVAB) [110]. A total of 1,618 doctors were recruited, but this only satisfied 10 percent of the observed shortage of doctors [111]. To address the lack of interest of Brazilian doctors to work in remote regions, the Brazilian government entered an agreement with the Pan American Health Organization (PAHO) and Cuba. This agreement resulted in thousands of Cuban doctors being sent to Brazil to fill the program vacancies. Within six months, the program introduced 11,430 doctors to these localities (79 percent Cuban), and within two years 18,240 doctors had been placed [108, 112]. According to the Ministry of Health, by the two year mark, Mais Medicos physicians were present in over 70 percent of Brazilian municipalities, including 34 indigenous health districts, and served 63 million people (Brazil has a population of about 210 million today).[2] Therefore, Mais Medicos led to a significant increase in the availability of health care providers in much of the country.

The program faced severe criticism from several sources. During the months leading up to the October 2013 vote that codified the program into law, many lawmakers, healthcare organizations, and citizens mainly protested two features of the program. First, one of PMM's goals, to restructure medical training at the University level, included adding two years to the physician field of study during which students would be sent to remote regions of the country to finish their degrees with practical hands-on training. This was perhaps the most controversial part of the program, and students around the country as well as university officials protested against this measure. The other main criticism was about the thousands of doctors from Cuba the government had agreed to hire. Although the government argued that without the Cuban doctors most of the positions created by PMM would remain unfilled, many protested what they perceived as foreign workers displacing Brazilians in government funded jobs. Additionally, the program dictated that Cuban doctors participating in this exchange program did not have to revalidate their medical degrees, bypassing the norm of physicians with foreign medical degrees having to take difficult

[2] http://maismedicos.gov.br/resultados-para-o-pais

examinations proctored by the Ministry of Health before practicing medicine in the country. The government had strong incentives to prevent Cuban exchange doctors from revalidating their degrees. If they were allowed to do so, under the Brazilian legal system, they would be free to practice medicine in any position of their choosing, meaning they could leave the positions assigned to them under PMM. However, the public criticized this feature of the program, arguing that the government would be placing under-qualified workers in these vacancies. For example, just 11 days after the announcement of PMM, the Federal Council of Medicine (CFM) filed a civil action against the government for the termination of the program, based on the criticisms described above. Later that month, students and doctors around the country went on strike to protest PMM, mostly about its intent to add two years to the medical training program [113]. Ultimately, the Brazilian legislative body voted in favor of PMM and officially codified the program into law in October of 2013 as law No. 12871.

3.3 Mais Medicos Implementation

Since its inception, PMM has allocated physicians to positions around the country through a series of open calls with three-year employment contracts. The Ministry of Health created an online site where trained physicians, foreign and domestic, could apply for one of the posted primary care physician vacancies, and municipalities could register to receive these physicians. In practice, almost no foreign physicians participated in PMM outside of the Cuba/PAHO agreement. The first call took place on July 8, 2013. A total of 3,511 out of 5,565 Brazilian municipalities registered, requesting 15,540 physicians. On August 5, 2013 the MoH published the final results from the first call: 1,753 physicians would be sent to 626 municipalities. The low registration of trained physicians prompted the Ministry of Health to seek foreign supply of physicians, as discussed below.

The remaining MoH official PMM calls (20 to date) followed a similar process. The MoH posted an open call during which physicians and municipalities had on average two weeks to register themselves. At the first peak of the program in January of 2016, there were 19,018 PMM physicians across the country. There was a second peak in October of 2017 with 19,068 doctors placed. As can be seen from the left panel of Figure 3.1, the expansion of the program was rapid. Within a year and half the program grew to almost 20,000 doctors per year and roughly remained at that level. Moreover, the right panel of Figure 3.1 shows that the program also expanded rapidly in terms of participating municipalities. The majority of municipalities that would receive PMM doctors received them within the first two years of the program.

Although physicians of any nationality were eligible to apply for jobs through PMM, applications were given differing priority levels. First preference was given to Brazilians trained in Brazil, then to Brazilians trained abroad, and lastly to foreigners trained abroad. Foreigners trained in Brazil were considered in the first priority category. This feature of the program resulted from the political debate about foreign doctors potentially displacing domestic doctors. All applicants were required to submit a statement indicating they had at least minimal knowledge of Portuguese, had no outstanding criminal issues, and that they were in good standing with their domestic military requirements if male, among other requirements. Only once applicants' documents and credentials were validated were they able to move on to the phase of enrollment where they could submit rankings for their four most preferred municipalities. If spouses or partners were also applying to the program, they were allowed to indicate preference for positions in the same municipalities. Once applicants were approved and placed in positions, their work permits were issued for three years, subject to renewal.

From the first open call it became clear that there existed a significant shortage of Brazilian physicians interested in participating in PMM. Although PMM offered competitive wages and benefits, the relocation outside of metropolitan regions required of participating physicians was likely an important deterrent. Nunes [114] uses administrative data of all Brazilian doctors who graduated from medical school between 2001 and 2013 and structurally estimates doctors' location choices to show that Brazilian doctors strongly prefer metropolitan regions for job location. Furthermore, although wages and health infrastructure are relevant, most predictive is where the doctor graduates from medical school, and medical schools are concentrated in metropolitan regions. As a consequence, the federal government entered a multilateral cooperation agreement with Pan American Health Organization / World Health Organization (PAHO/WHO) and Cuba for Cuban doctors to relocate to Brazil and fill the vacancies in PMM.

The Ministry of Health determines the number of vacancies allocated to registered municipalities using complex rules. Table 3.1 shows the seven mutually exclusive categories defined by the PMM program rules. One component of the categorization is the PAB (*Piso da Atenção Básica Fixo*) is the Basic Care Floor that consists of a fixed and variable amount of funds the federal government transfers yearly from the National Health Fund to municipal health funds to finance health care services. This amount is determined through the construction of a score from 1 to 10 that incorporates municipal measures of GDP per capita, percent of population with health insurance, percent of population receiving Bolsa Familia, percent of population in extreme poverty, and demographic density estimates. The PAB has four groups where Group I receives the highest transfer and has the smallest population and group IV receives the lowest transfer. Transfers are comprised of a fixed and a variable amount per inhabitant. PMM uses information on the fixed component of the PAB to determine program eligibility category. Information on the precise construction of the score is not publicly available. However, Table 3.2 presents the definitions of the PAB groups used to construct the PMM municipality profiles as of July 2013. As is evident from Table 3.1, most municipalities fall under Profiles 4 and 7, as do the allocation of doctors per capita. As expected, most birth records in the sample fall under Profile 3 municipalities due to the population size of capital cities. From Table 3.1, we can see there was active targeting of doctor placement under the PMM program towards municipalities with lower PAB scores, smaller populations, and in extreme poverty. Furthermore, tabulations from Girardi [115] confirm that the Northern, Northeastern, and small municipalities had the highest scarcity levels of doctors prior to 2013 and saw large increases in coverage from Mais Medicos doctors compared to other regions.

3.4 Literature Review

Although the literature on the effects of medical care on health outcomes is vast, there exist few studies analyzing the effect of physician access on birth outcomes. Instead, several studies focus on the relationship between physician supply or quality and outcomes such as adult mortality, demand for medical services, nurse supply, reported quality of care, HIV transmission, among other outcomes [116–123]. For birth outcomes, much of the literature explores the causal link between quality of care, such as prenatal care, and birth outcomes, not the effect of physician supply itself [124–130]. There are some studies explicitly analyzing skilled medical professional supply and birth outcomes. Okeke et al. [131] examines the effect of skilled birth attendant instead of a traditional midwife present at the birth on delivery complications but finds no significant relationship. Macinko, Starfield and Shi [116] finds that increased supply of primary care physicians is associated with improved birth weight in the US, however, a study from the epidemiological literature Shi et al. [132] finds a null effect of increased physician supply on birth weight in the US. Furthermore, an analysis in the public health literature finds an additional physician per 10,000 people in China between 2012 and 2017 decreased low birth weight incidence, albeit by 0.106 percent [133]. How physician supply affects birth outcomes remains understudied, particularly in the field of economics.

As PMM is a unique and large-scale program, there is a growing literature examining its impacts on a range of health outcomes. Fontes, Conceição and Jacinto; Hone et al; Mattos and Mazetto; Özçelik et al; Carrillo and Feres [134–138] estimate the program's impact on hospitalizations, hospitalizations for ambulatory care sensitive conditions, hospitalizations for primary care sensitive cardiovascular conditions, amenable (preventable) adult mortality, health visits with specialists, physician-provided prenatal care, and birth outcomes such as gestational length and infant mortality. The results present a lukewarm assessment of the success of PMM in improving health outcomes. For example, Carillo and Feres [138] find no effect of PMM on prenatal care visits and infant mortality, however the authors do not address the data collection changes that occurred in 2011 that the Brazilian Ministry of Health report affected measures of gestational length and prenatal visits. Additionally, while Hone et al. [135] find a reduction in amenable adult mortality, Mattos and Mazetto [136] find no effect on adult mortality. However, Özçelik et al. [137] find a decrease in hospitalizations, though only starting four years after the start of PMM. This reduction was only present in a small subset of hospitalizations, those related to primary care sensitive cardiovascular conditions. This analysis contributes to the literature on the relationship between physician supply and birth outcomes and that of the PMM program by focusing on the effects of PMM on birth weight and incidence of low birth weight.

3.5 Data

The Sistema de Informações de Nascidos Vivos, or vital statistics data (SINASC), provides information from the birth certificates of every registered birth in Brazil. The data include about 300,000 births per year between 1996 and 2018, with 97.80 percent recorded as occurring in a health institution. While it is not clear how many births that occur outside of health facilities remain unregistered, much of the literature consider SINASC nearly universal and the Ministry of Health reports 100 percent coverage by 2011 [139, 140]. The publicly available data provides the municipality where the birth occurs and the municipality where the mother resides. For this analysis, the municipality where the mother resides will be the primary geographical identifier as women mostly seek medical care where they live. In addition to birth weight, type of birth (singleton, twins, or more), and gestational length, SINASC also provides rich demographic information on the child and mother. Mothers' characteristics include municipality of birth, race, age, educational attainment, marital status, and other information about number of past live and deceased births, past c-sections, and date of last menstrual cycle. While the distribution of these demographic characteristics could be impacted by PMM, for example, by changing the composition of mothers, and are therefore not included as controls in the regression models presented below, considering how these compositional changes may have materialized is of interest. Finally, SINASC also provides demographic characteristics of the newborn and information about the pregnancy such as newborn's race, gestational length of the pregnancy, number of prenatal visits, the month of pregnancy at first prenatal visit, fetal presentation at birth, and other details.

Beginning in 2009, the Ministry of Health (MoH) undertook a restructuring of the SINASC data with the stated of goal of better capturing the inequalities in access to health care by different racial groups. After pilot testing across a small sample of municipalities, the new SINASC survey launched in 2010. To ensure the continuity of a historical series, any new formulations of original survey questions were recorded as new values in the data, while still recording answers to the original survey questions. For example, number of prenatal visits in the old survey version was recorded as a categorical variable in intervals of about 5 weeks. In the new survey, the precise number of visits is recorded as a new variable, but it is also recorded in the original categorical variable constructed from the continuous survey question. Municipalities only gradually rolled out the new survey throughout 2010 until all received it by January 2011. However, municipalities had differential dates of adoption of the new survey such that the 2011 data are comprised of 42 percent old surveys and 58 percent new surveys, with the southeast region having a lower adoption rate than the northern region. [140]. The birth weight and birth date variables used in the analysis below are not part of the updated variables. However, since variables such as gestational length and prenatal visits change with the updated surveys, using them as additional outcomes prior to 2011, in the same model, would yield estimates that represent the changes in the data collection method instead of changes from treatment. For further details on how survey questions changed in 2011 see the Data Appendix. From its report on the data consolidation, the Ministry of Health [140] calculates that coverage of all live births increased from 97 percent to 100 percent between 2010 and 2011, calculated as actual reported births to projected live births by IBGE. The MoH reports that the only variables that experienced any meaningful deviation from historical trends were gestational length and race of mother. However, MoH argues these changes do not reflect a change in the composition but rather an improvement in the quality of the data. This is also discussed in more detail in the Data Appendix.

Information about the program's allocation of doctors was obtained through a request to The Ministry of Health. This information includes the number of PMM doctors in each municipality on a monthly basis, as well as the PMM profile of each municipality. Brazil has 5,565 municipalities and they have full autonomy in the hiring of healthcare professionals in the public sector through the Unified Health System (SUS). Data from several government agencies are included as time-varying municipality-level controls. IBGE (Brazilian Institute of Geography and Statistics) provides yearly municipal GDP and population estimates, while the Ministry of Finance provides yearly municipal total spending, and expenditures on welfare, judicial, and security spending. The Ministry of Citizenry provides the number and amount of Bolsa Familia payments by municipality. Bolsa Familia is a large cash-transfer program designed to increase nutrition and schooling for poor children. The IBGE also provides socio-demographic information from the 2010 Demographic Census including population density, municipal Human Development Index (MHDI), literacy rates, and the share of population that is Black. Number of primary care physicians in each municipality at baseline (June 2013) is obtained from Cadastro Nacional de Estabelecimentos de Saúde (CNES), an administrative data record of physicians and other health infrastructure. Primary care physicians include those in family health, clinical medicine and pediatric medicine following Girardi et al. [115]. Municipality-level characteristics are included to control for time-varying attributes of municipalities that could affect birth outcomes outside of PMM. For example, we may expect municipalities that have higher levels of spending to exhibit better birth outcomes as they potentially invest more in the health and wellness of their residents. Therefore, fluctuations in the levels of municipal spending may be an important explanatory variable.

Table 3.3 presents summary statistics from the birth records data for births occurring between January 2006 and December 2017, a total of 144 months. Due to the data changes in 2011, the values for gestation, the type of birth attendant, whether the labor was induced, number of prenatal visits, and the number of past vaginal and cesarean deliveries are from a sample starting in January 2011. Note that average birth weights throughout the sample period are relatively healthy. The average is 3,188 grams, while the average in the US is 3,389g (for single births) [141]. Brazil is above the average for OECD countries on incidence of low birth weight (less than 2,500 grams [142]) at 8.7 percent, but much lower than many low-income countries. Sixty-five percent of births are attended by a physician, and less than 1 percent are attended by a midwife. Average gestation is about 38.5 weeks, and labor is induced 20 percent of the time. Average prenatal visits are also at healthy levels, at about 8. For comparison, the average for the US in a similar time period is 10. We can see that most women make their first prenatal care visit in the second to third month of their pregnancy. The sample is comprised almost entirely of white and mixed-race newborns, with 4 percent Black and less than 1 percent Indigenous. Finally, cesarean births are very high in Brazil. In this sample, over half of all deliveries are by cesarean-section. This is high even compared to the U.S. where in 2013 the cesarean-section rate was also high at about 33 percent, but even more so when compared to the 2013 OECD average of 25 percent [143, 144].

Table 3.3 also presents characteristics of mothers. Mean age at birth is almost 26 year old, and the average woman has at least one previous pregnancy. Delivery was more likely to be vaginal than cesarean, reflecting that the shift towards cesarean-sections has been a recent development. Most mothers have at least some high school education, with only 1 percent of mothers with no formal education. Finally, only 2 percent of the sample of mothers reside in rural and remote areas. These descriptive statistics indicate healthy average birthweights for newborns, possibly suggesting a limited scope for PMM to affect this outcome. The total sample from 2006 to 2017 includes over 34 million birth records, while the sample from 2011 to 2017 includes over 20 million records. Mothers in the sample are most likely observed more than once, for each birth. However, the publicly available data does not uniquely identify them, so it is not possible to track them over time, over multiple births. Requesting the restricted-access SINASC records may be of interest for further analysis.

Figure 3.2 shows the frequency distribution of the birth weight records for the full sample of years 2006-2017 and the subsample of years 2011-2017. We can see from these figures that the distribution of birth weight changes between the two samples, with a discrete jump in both at the 3,100g value, but at 3,200g for the later sample. While it is common for birth weight distributions to exhibit bunching at the 500 gram intervals when hospitals do not use precise scales for the measurements, this sample does not exhibit this behavior [145–147]. Additionally, bunching may appear most often at the low birth weight cutoff of 2,500g, and we also do not see evidence of that in this sample. While the records to do not follow a normal distribution precisely, this provides suggestive evidence that the birth records are at least very precise with the birth weight measurements in Brazil.

3.6 Methodology

I employ three approaches to identify the causal effect of PMM on birth outcomes. The first specification uses differences-in-differences (DiD) with a binary treatment measure. By exploiting the timing and geographical variation in the implementation of PMM, a DiD approach will estimate the causal effect of PMM on birth weight outcomes, conditional on parallel pre-trends in the outcome variable between the treated and control group. As discussed below, we reject the parallel pre-trends assumption in this sample. Nevertheless, to better understand what drives the variation in this estimator, the second specification employs a DiD with a continuous treatment. In this specification, the estimated treatment effect is informative about the importance of the intensity of treatment, defined as the number of PMM physicians per 10,000 residents in a given municipality. It could be the case that PMM was only effective in improving birth outcomes in the municipalities that were more intensely treated. This estimate is causal if parallel pre-trends cannot be rejected and if we can assume that going from 0 to k treatment units is stable over time. Note the additional assumption needed to causally identify this parameter compared to the binary treatment DiD. As with the first DiD specification, the parallel pre-trends test is rejected. Finally, given the results from the parallel pre-trends tests for both DiD specifications, and the subsequent concerns we may have about the causality of those estimates, an instrumental variables approach uses the program eligibility rules to instrument for number of PMM physicians in municipalities. If the instrument is valid, by satisfying the exclusion and exogeneity assumptions discussed below, then this 2SLS approach will yield a causal estimate of the Local Average Treatment Effect (LATE). This estimate will be informative on the treatment effect of PMM for those who are affected by the instrument, or the compliers. Unlike the DiD models, we do not need parallel pre-trends in the outcome of interest to causally identify this parameter.

Comparing the DiD estimates to the 2SLS estimates is instructive to shed light on the direction of the potential bias in the DiD specifications. Additionally, in the presence of heterogeneous treatment effects, which are quite plausible in this context (e.g. impact of an increase in PMM physicians in the public sector may affect mothers in different socioeconomic statuses differently), the DiD estimates of the average treatment effect (ATE) will not equal the LATE even if both estimates are unbiased. Furthermore, although PMM was implemented in most municipalities across Brazil, it likely did not affect all mothers, but rather the subset of mothers in the sample who were more likely to interact with public sector physicians for their medical needs. Therefore, we could argue that the LATE is a more appropriate estimate of the effect of PMM on birth weight outcomes over an ATE. Nevertheless, as will be presented in Section VII, all three of these models in the fully saturated specifications indicate no effect of increased physician supply on birth weight.

3.6.1 DiD: Binary Treatment

The first approach compares births in municipalities with PMM physicians to births occurring in municipalities without PMM physicians:

$$\mathrm{Y}_{imt} = \alpha + \beta PMM_{mt} + \delta_t + \delta_m + \Gamma X_{mt} + \epsilon_{imt}$$

where Y_{imt} is the outcome of interest for birth *i*, in municipality *m* and month *t*. PMM_{mt} is a binary indicator equal to 1 for municipalities who ever received PMM doctors from July 2013 and on. Therefore, treatment ”turns on” and does not turn off for municipalities. In practice, treatment did not turn on at the same time for every municipality. However, as discussed previously, the program expanded rapidly to all treated municipalities within two years. I explore the implications of this variation in treatment timing in Section VIII. X_{mt} is a set of time-varying municipality controls that include log GDP, log population size, log share of total municipality spending on welfare, log share of municipal spending of municipal GDP, the share of residents that receive Bolsa Família and the average Bolsa Família transfer per recipient. All of these municipal controls vary at the month level, except for population which varies at the yearly level. In more saturated models, instead of including these time-varying municipal controls, I include their values at baseline in 2013 interacted with a linear time trend and state-specific time trends [148]. These fully saturated specifications serve as robustness checks. Finally, δ_m and δ_t are municipality and time fixed effects, and ϵ_{imt} is the idiosyncratic error term.

The identifying variation in this model is at the municipality by time level, given that assignment to treatment and control groups is defined by a municipality registering for and obtaining physicians from PMM. We expect β to recover the causal effect of easing doctor supply constraints through PMM on birth outcomes if treated and control municipalities would have experienced similar changes in birth weight trends without PMM. Thus, the identification assumption is that in the absence of PMM, newborns in treated and control municipalities would have experienced similar changes in birth outcomes over time, and that the timing was not endogenous to changes in the time series. The main threat to identification in this model and the subsequent models lies in the fact that PMM did not randomly assign physicians to municipalities. On the contrary, government reports indicate the MoH designated municipalities with worse socioeconomic conditions with higher priority physician allocation. Physician allocation is further discussed in the next section. However, the set of municipal controls are included to attempt to address this endogeneity. This set of socioeconomic variables are included since they may be simultaneously affecting the implementation of PMM and birth outcomes. Additionally, the wide set of municipal and state interactions with linear time trends are included to address potential time-varying unobservables that could be affecting birth outcomes differentially between the treated and control municipalities.

Figure 3.3 shows mean birth weight trends over the sample period by treatment status, for the full sample as well as the rural sample. While both groups experience increasing birth weight trends post-2013, the treatment group begins to experience this improvement in trends in 2011. This raises concerns about the increase in survey coverage from 2011 potentially having an effect on the observed composition of births, particularly because the increase in coverage was driven by municipalities with lower socioeconomic conditions, and these are exactly the municipalities targeted for PMM treatment. We do not see a similar change in 2011 from the rural sample. The change in survey coverage may be a driver for the differences we observe in birth weight trends between treatment and control groups, although the MoH report indicates no significant compositional change. If true, then the current results could be negatively biased by this shift in the sample composition if the increased coverage does not occur for the same types of regions across treated and untreated municipalities.

Parallel pre-trends is formally tested for the pre-treatment years 2006-2012 in Table 3.4 which presents results from testing the joint significance of the β_t:'s from the following model:

$$\text{Y}_{imt} = \alpha + \sum_t \beta_t PMM_{mt} * Year_t + \delta_m + \delta_t + \Gamma X_{mt} + \epsilon_{imt}$$

The results of this parallel pre-trends analysis raises concerns about the validity of the assumption for this model. In general, the rural sample estimates are noisier than those in the full sample. For each sample, the first columns shows estimates with time varying municipal controls, the second instead includes their baseline values interacted with a linear time trend, and the third column includes state-specific linear time trends, for additional robustness checks. For the full sample, we reject parallel pre-trends in both birth weight in grams and incidence of low birth weight in the more saturated models. There is the concern that given the large sample size, even a small difference from zero will be statistically significant. Yet, the magnitudes of the estimates are economically meaningful. For example, a treatment effect of 11 grams would be substantial given the current literature on birth weight outcomes. From this analysis, we should be cautious in interpreting the estimated coefficients from this model as causal.

3.6.2 DiD: Continuous Treatment

The second approach exploits the variation in the number of PMM physicians allocated throughout municipalities over time to learn about the effects of treatment intensity. Treatment is defined as the number of PMM doctors a municipality has in a given month per 10,000 residents. Adjusting the counts of physicians allocated by the program for population addresses the relationship between the population size of a municipality and its demand for physicians. The DiD model with a continuous treatment variable is as follows:

$$\mathrm{Y}_{imt} = \alpha + \beta(PMM\ doctors)_{mt} + \delta_t + \delta_m + \Gamma X_{mt} + \epsilon_{imt}$$

Where Y_{imt}, X_{mt}, δ_m and δ_t remain as in the previous model. $(PMM\ doctors)_{mt}$ now measures number of PMM doctors per capita in a municipality on a given month using yearly municipal population estimates.

In this continuous treatment, or ”dosage” DiD model, the identifying variation is generated through the variation in number of PMM physicians over time and municipalities. Therefore, the model estimates the effect of an additional doctor (per 10,000 residents) on birth outcomes. In other words, it is comparing a child born in a municipality with a smaller allocation of PMM physicians to a child born in a municipality with a higher allocation of PMM physicians. This model faces similar endogeneity concerns as the binary treatment DiD model. Given the ambiguity of how the MoH decided how many physicians to allocate across the municipalities throughout the open calls, there likely exist unobservable factors driving the presence of more or less PMM physicians across municipalities.

To gain more insight on the placement of PMM physicians, Figure 3.4 and Table 3.5 show some suggestive evidence on the targeting of PMM assignment. Figure 3.4 plots the number of PMM physicians in each municipality during the middle of the program, January 2016, against the number of primary care physicians present at baseline, June 2013. There is a positive relationship, as expected, since bigger municipalities like São Paulo and Rio de Janeiro have higher physician demand. The right panel of Figure 3.4 shows this relationship adjusting for population size by using the number of physicians per 10,000 residents. The apparent rays in the scatter plot result from the fact that the number of PMM physicians is most often a small whole number and population is not changing substantially. Both of these figures look almost identical when dropping São Paulo, Rio de Janeiro, and Fortaleza, except with the left graph having a steeper line of fit. If there existed perfect program targeting, we would expect a distinct negative relationship between PMM physicians and baseline primary care physicians. As described in Section II, the MoH reports to have assigned municipalities physicians using a measure of physician shortage. Therefore, we should expect that the lower the baseline level of primary care physicians, the more PMM physicians a municipality would be assigned. However, we observe a weak negative relationship at best. This suggests imperfect targeting of PMM physician allocation.

Table 3.5 presents more information on physician allocation. The coefficients are correlations between an array of municipal characteristics and the number of PMM physicians present in municipalities in January 2016 per 10,000 residents. From the most saturated model with region fixed effects shown in column (9) we can see that the Black share of the population, the municipal Human Development Index of income (MHDI), and the baseline level of primary care physicians are statistically significant predictors of the number PMM physicians. A higher Black share of the municipal population is associated with fewer PMM physicians, as is a higher MHDI in income. Meanwhile, a higher stock of primary care physicians at baseline is associated with more PMM doctors, as well as being a municipality in the Midwest or South compared to Southeast (where São Paulo and Rio de Janeiro are located). Again, if there were perfect program targeting, we would expect a negative relationship between baseline number of primary care doctors and PMM doctors. The strong negative relationship between the income MHDI and the number of PMM physicians suggests targeting was not random.

Figure 3.5 plots mean birth weight and mean incidence of low birth weight across municipalities with no PMM doctors, those with PMM doctors per 10,000 below the mean of 1.77, and those with more than the mean. This graphical representation of the pre-trends in birth outcomes implies the parallel trends assumption prior to PMM may not hold. In both cases, it seems that the high exposure group has more similar changes in birth outcomes over time to the control group than the low exposure group. It is also interesting to note that the high exposure group has a lower incidence of low birth weight and correspondingly, higher average birth weight. This is also suggestive evidence of imperfect targeting, given the groups were defined by population-adjusted number of PMM physicians. These trends are tested formally in Tables 3.6 and 3.7. Each panel of the two tables tests the exposure groups separately, where Table 3.6 shows the results for the full sample and Table 3.7 shows them for the rural sample. We can again reject parallel pre-trends for all specifications in the full sample, though we cannot reject them in most specifications in the rural sample. As with the trends presented for the binary treatment model, the magnitudes of the coefficients are economically meaningful. However, it seems it is one or two years driving the statistically significant difference. Nevertheless, estimates in both tables present models with the wide set of municipal characteristics interacted with linear time trends, suggesting we are still failing to account for the source of the differential trends in outcomes pre-treatment.

3.6.3 Instrumental Variables

The third and final specification uses an instrumental variables approach. As shown in Tables 3.1 and 3.2, the MoH defined particular ”rules” for physician allocation such that certain municipality profiles received significantly more physicians than others. As discussed previously, municipalities in profiles 4 and 7 were assigned many more physicians per capita than the other 5 profiles. Again, some of these profiles were defined using the municipality's PAB group, ranging from I to IV. To use these program rules as instruments for physician placement, they must be relevant and accurately predict doctor placement, but must also satisfy the exclusion restriction and only affect birth outcomes through PMM physician placement. These profiles were not reported as used in any other way by the MoH. They were initially defined in the legislation announcing the PMM program, and only included in documentation about PMM. On the other hand, only using the PAB group definitions as instruments would indeed violate the exclusion restriction. The amount of funding a municipality receives yearly from the federal government for health care spending is determined by which PAB group it falls under. Consequently, these definitions could have direct effects on birth outcomes. Therefore, the excluded instruments are indicators for a municipality's PMM profile for those profiles not defined by PAB groups [3]. Observations from municipalities classified on the basis of PAB are dropped from the sample.

The first and second stage equations for this model are:

PMM Doctors$_{mt} = \eta + \phi \sum_1^k Profile_{mt} + \Gamma X_{mt} + \epsilon_{imt}$

[3] IV models controlling for PAB funding yield very similar estimates.

$$\mathrm{Y}_{imt} = \zeta + \psi(PMM\ \widehat{Doctors})_{mt} + \Gamma X_{mt} + \epsilon_{imt}$$

where ΓX_{mt} and ϵ_{imt} remain as in the previous models.

3.7 Results

Tables 3.8 and 3.15 present the parameter estimates from the different methodologies for the fully saturated specifications, for the full and rural samples respectively. Tables 3.10, 3.11, 3.12, and 3.16 present specifications for each model. For Tables 3.10, 3.11, 3.12, and 3.16 column 1 presents estimates from a specification with no covariates, column 2 includes yearly municipal covariates, column 3 includes baseline municipal covariates (measured in 2013) interacted with a linear time trend, and the fourth columns also includes state-specific time trends.

For both DiD approaches we observe that as we add covariates and time trends to account for differential trends across treated and control municipalities, the point estimates become statistically insignificant for the full sample. The fact that the point estimates change significantly as we include the sets of interactions confirms the presence of differential trends between treated and control municipalities that bias the ATE away from zero. For the rural sample, none of the DiD specifications yield statistically significant effects, as seen in columns 5 - 8 in Tables 3.10 and 3.11. In the binary treatment DiD specification in Table 3.10, column 4 suggests that births occurring in municipalities with higher physician supply due to PMM saw an average decrease in birth weight of 0.425g, from a base of 3,232g. In Table 3.11, from the continuous treatment DID we see an estimated effect of 0.088 x 1.77 = 0.155g in column 4. However, it is positive, rather than negative as in the binary DiD specification. If unbiased, this could suggest that treatment intensity is an important dimension for estimating the effect of PMM on birth weight, perhaps indicative that only mothers in higher-intensely treated municipalities experiences gains in their newborns' birth weight. However, neither estimate is significantly different from zero. The corresponding estimate for the continuous DiD approach in the rural sample in column 8 is larger than in the full sample, as we may expect if we believe PMM was most valuable in these areas, at 0.435 x 1.77 = 0.770g, though a noisy estimate.

Finally, Tables 3.12 and 3.16 present the results from the naive OLS and instrumental variables approaches. The OLS estimates of a regression of birth weight on the number of PMM physicians per capita in Table 3.12 suggest, unsurprisingly, a strong positive relationship between number of physicians per 10,000 residents and birth weight and incidence of low birth weight, though not statistically different from zero in the fully saturated specification. If unbiased, OLS also estimates the ATE. To learn about the direction of the bias of the OLS estimates, we can compare them to the 2SLS estimates from Table 3.16 which presents the effects of PMM physicians on the compliers, those mothers who would seek medical care from PMM physicians if their municipality was eligible, and would not otherwise. If we consider these LATEs unbiased, they suggest the OLS estimates are considerably downwards biased. This may be explained by the fact that the PMM implementation strategy placed more PMM physicians per 10,000 residents in the more socioeconomically disadvantaged municipalities, which have worse birth weight outcomes pre-intervention, and are more metropolitan. Rural municipalities in Brazil have higher mean mean birth weight levels. Many studies have shown this paradoxical relationship between higher mean birth weight coupled with higher incidence of low birth weight in rural Brazilian municipalities compared to urban ones [149, 150]. Note the first-stage F statistic fluctuates significantly across specifications in columns 1 - 4 of Table 3.16, however it is high even in the fully saturated model in column 4 at 33.43. Nevertheless, the LATE estimates are implausibly large at 44g. There is no other example in the literature measuring such a large effect on birth weight from any interventions or natural disasters. Columns 5 - 8 present the rural sample results where the LATE estimates fluctuate substantially across specifications, however none of the estimates are statistically different from zero. These results could be the result of weak instrument issue that yields upward-biased IV estimates. This is discussed below in section VII. The first stage results are presented in Table 3.13.

3.8 Robustness Checks

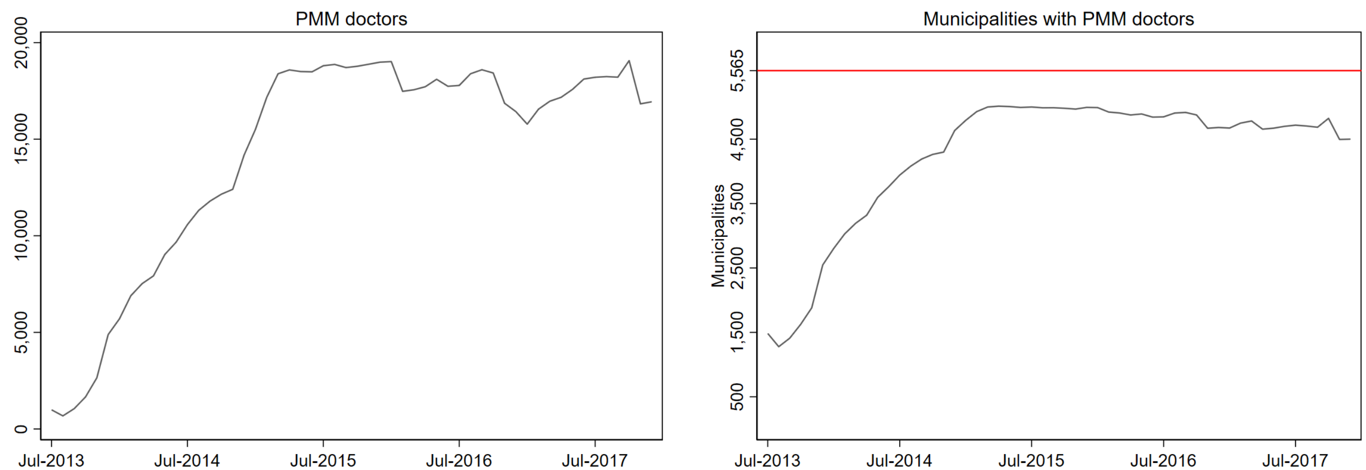
As shown in Figure 3.1, all municipalities did not begin receiving PMM physicians at the same time. It took the MoH almost 24 months to fully implement PMM in all eligible municipalities. Consequently, we may be concerned about the interpretation of the ATE estimate from the binary treatment model in Table 3.10 in the presence of this variation in treatment timing. To address this concern, Table 3.14 presents ATE estimates that incorporate the variation in PMM implementation timing across municipalities. In this specification, observations are considered treated if the birth occurred in a municipality that had PMM physicians present in that corresponding month. This definition of treatment allows for treatment to ”turn off”. If unbiased, these estimates should present an ATE estimate that does not suffer from the attenuation bias arising from the mis-classification of municipalities as treated in months where in fact, they are not. We see from column 3 of Table 3.8 a positive and statistically significant ATE estimate for the effect of PMM on birth weight. The point estimate of 2.663g is significant at the 2 percent confidence level in this fully saturated specification. The rural estimate in Table 3.15 is negative and larger in magnitude, but not statistically significant. We can interpret this estimate as suggesting a positive impact from the PMM physician expansion on birth weight.

It is also possible that the IV estimates in Table 3.16 suffer from a weak instrument problem. Recent work by Lee [151] has shown that we should no longer assume that an F-statistic greater than 10 implies a strong first stage, but instead should consider the threshold to be 104.7. Given the results in Table 3.16, it is possible that the IV estimates suffer from this concern. To address the potential weak instrument problem, I estimate the instrumental variables specification using Limited Information Maximum Likelihood (LIML), which yields test statistics robust to weak instruments. These results are presented in Table 3.17. The IV estimate from Table 3.16 and the LIML estimate from Table 3.17 are practically identical in all specifications in both the full and rural samples. For example, the LATE estimate from 2SLS yields a point estimate of 44.180g while LIML yields an estimate of 44.534g. Both estimates are statistically significant at the 5 percent level. This suggests the IV estimates do not suffer from a weak instrument problem, and the estimated effect magnitudes are indeed quite large compared to the literature.

3.9 Conclusion

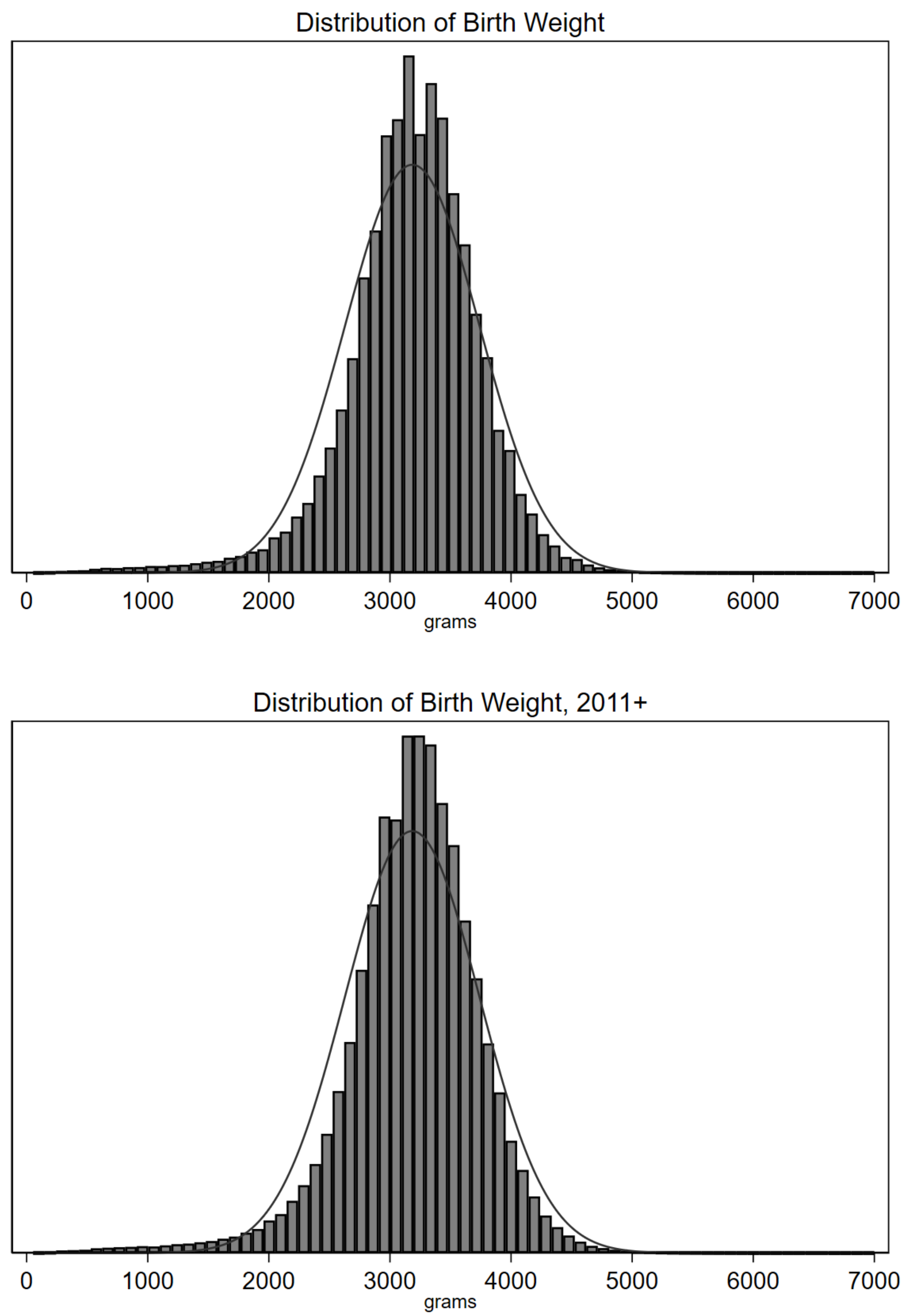
This project attempts to contribute to the literature on improving health outcomes by analyzing the effect of introducing skilled medical professionals in areas where they are in dangerously low supply. Interventions of this kind have a large margin of potential impact and speak to a common problem throughout the developing world, the shortage of appropriate medical care for certain groups of people. Quantifying the effects of such a policy as the Mais Medicos Program can serve to shed light on the path towards improving the general well-being of a significant portion of the world's citizens. The IV estimates, though likely biased upwards, suggest that PMM resulted in higher average birth weight for children throughout Brazil. However, I find no improvement on the incidence of low birth weight or any weight effects for those living in rural parts of the country. Hence, these results imply PMM did not affect the most vulnerable mothers and their pregnancies.

Figure 3.1: PMM Doctors



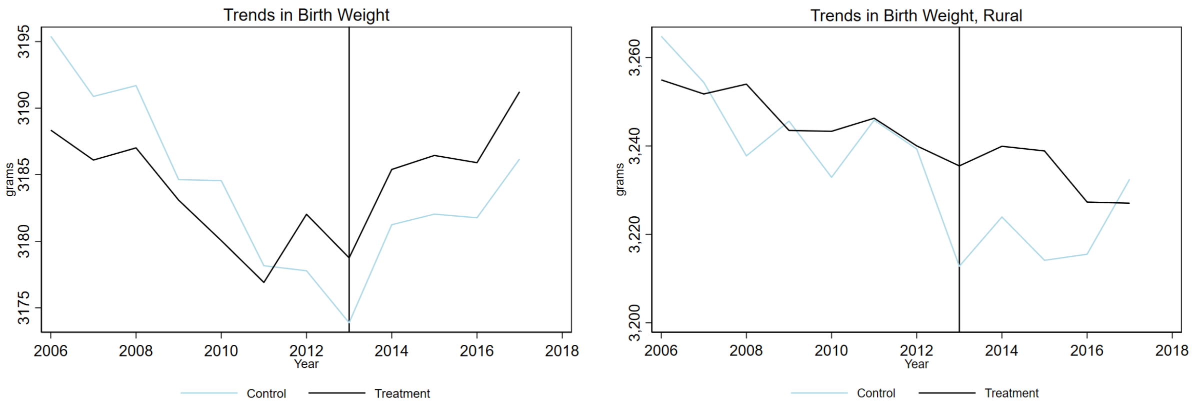
Note: Data from Brazil’s Ministry of Health. The figure on the left presents the monthly number of PMM doctors across all municipalities. The figure on the right presents the monthly number of municipalities with PMM doctors. Total Brazilian municipalities is 5,565.

Figure 3.2: Distribution of Birthweight in Brazil



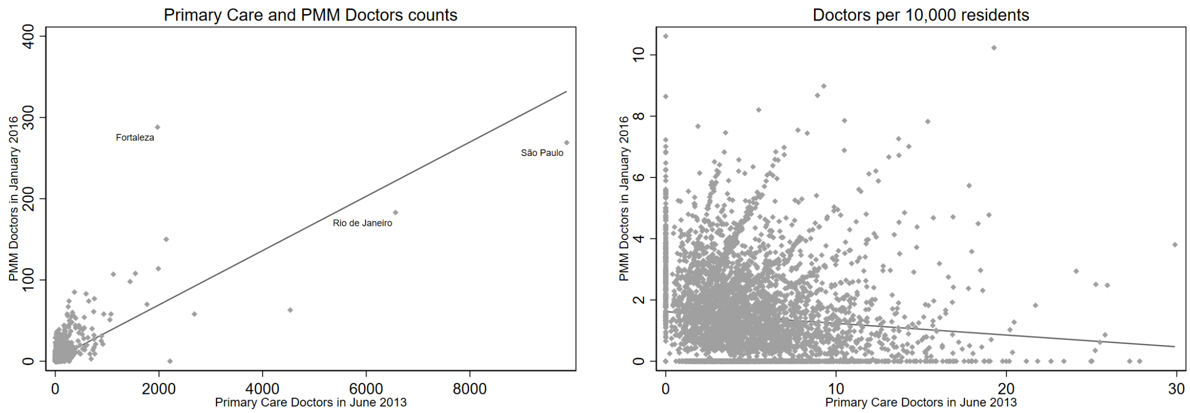
Note: Data from SINASC. The figure on the top presents the distribution of birth weight in grams for the 2006-2017 sample. The figure on the bottom presents this distribution for the 2011-2017 sample.

Figure 3.3: Trends in Birthweight



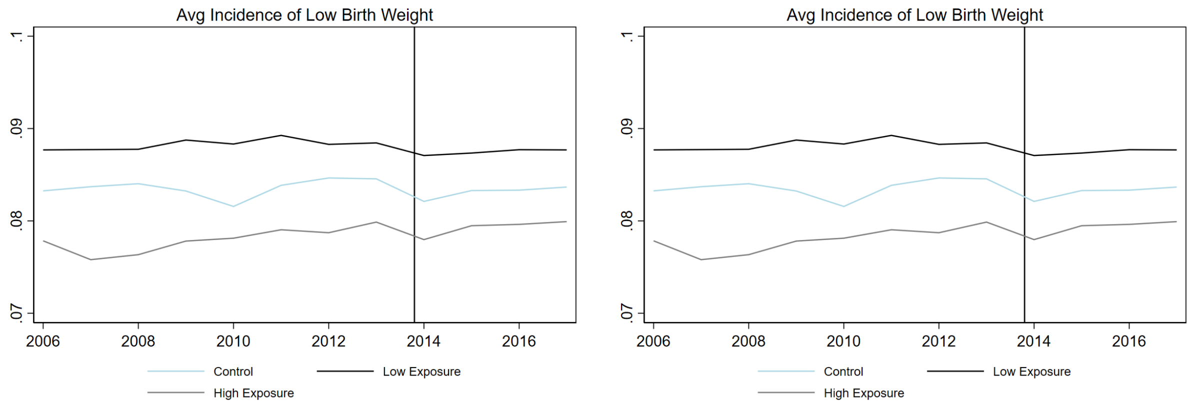
Note: Data from SINASC.

Figure 3.4: Primary Care and PMM Doctors



Note: Data from SINASC, IBGE, and Ministry of Health. The graph in the left panel plots the number of PMM doctors present in a municipality in January 2016 against the baseline number of primary care physicians in June 2013. The right panel adjusts the measures of doctors to per 10,000 residents.

Figure 3.5: Trends in Incidence of Low Birthweight



Note: Data from SINASC.

Table 3.1: Municipality PMM Profiles

Profile	Definition	Mun.	PMM Mun.	PMM doctors	Docs per cap	Births
Profile 1	PAB groups III and IV	277	240	3,627	189	3,934,137
Profile 2	PAB group II	775	596	3,715	745	2,516,988
Profile 3	Capital	536	501	12,835	624	14,431,061
Profile 4	PAB group I	1,622	1,173	4,714	2,007	2,473,041
Profile 5	G100	98	98	4,567	87	3,634,931
Profile 6	Vulnerable Area	554	493	4,402	761	2,308,578
Profile 7	Extreme Poverty	1,708	1,543	13,384	2,700	5,717,187
Total		5,570	4,644	47,244		35,015,923

Data comes from the Ministry of Health and SINASC. Municipality profiles were defined by the Ministry of Health and are mutually exclusive. Doctors per capita measures the number of doctors at the first peak of the program in January 2015 per 10,000 residents in baseline 2013 population estimates.

Table 3.2: PAB Definitions, 2013

Group	Definition
Group I	Municipalities with a score lower than 5.3 and a population of up to 50,000.
Group II	Municipalities with scores between 5.3 and 5.8 and a population of up to 100,000; and municipalities with a score lower than 5.3 and a population between 50,000-100,000.
Group III	Municipalities with scores between 5.8 and 6.1 and a population of up to 500,000; and the municipalities with scores below 5.8 and population between 100,000 - 500,000.
Group IV	Municipalities not included in the previous items.

Information comes from the Ministry of Health. The PAB is the yearly funds the federal government allocates municipal health funds from the National Health Fund.

Table 3.3: Descriptive Statistics - 2006 to 2017

Birth Outcomes		***Mother's Characteristics***	
Birth weight	3,184	Age	25.96
Low birth weight	0.087	Single	0.498
Very low birth weight	0.013	No education	0.011
Attended by physician	0.651	Yrs of ed: 1-3	0.052
Attended by nurse	0.047	Yrs of ed: 4-7	0.236
Attended by midwife	0.006	Yrs of ed: 8-12	0.513
Weeks gestation	38.48	Yrs of ed: 12+	0.169
Induced Labor	0.190	# of alive children	1.094
		# of deceased children	0.190
Newborn characteristics		# of previous pregnancies	1.250
Female	0.488	# of previous vaginal deliveries	0.761
White	0.402	# of previous cesarean deliveries	0.343
Black	0.036	Lives in rural area	0.021
Asian	0.003	Births	34,684,853
Mixed	0.504	Births 2010+	20,312,142
Indigenous	0.007		
Pregnancy characteristics			
C-section	0.529		
Multiple birth	0.020		
# Prenatal Visits	7.633		
Month of first prenatal visit	2.658		

Notes: Data from SINASC. Births in the sample consists of births between January 2006 and December 2017. Low birth weight is defined by WHO as birth weight under 2,500g, and low birth weight as those under 1,500g. Gestation, birth attendant, labor induction, prenatal visits, and the number of vaginal and cesarean deliveries are measured from 2010, not 2006.

Table 3.4: Test of parallel pre-trends for Binary Treatment DiD

	Birth weight (grams)						Low birth weight incidence					
	Full Sample			Rural Sample			Full Sample			Rural Sample		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
2007	1.631	-0.952	-0.968	9.176	-1.247	-4.125	-0.0006	-0.0003	-0.0002	-0.0037	-0.0003	0.0014
	(2.625)	(2.742)	(2.765)	11.57	(12.224)	(12.527)	0.0013	(0.0014)	(0.0014)	0.0057	(0.0060)	(0.0062)
2008	1.669	-2.754	-1.948	30.766	26.117	30.337	-0.0009	-0.0005	-0.0008	-0.0156	-0.0159	-0.0181
	(2.618)	(2.737)	(2.759)	11.543	(12.205)	(12.512)	0.0013	(0.0014)	(0.0014)	0.0057	(0.0060)	(0.0062)
2009	5.265	1.614	1.963	10.800	-0.574	0.540	0.0010	0.0006	0.0009	-0.0066	-0.0043	-0.0042
	(2.637)	(2.756)	(2.779)	11.644	(12.313)	(12.631)	0.0014	(0.0014)	(0.0014)	0.0053	(0.0061)	(0.0062)
2010	2.349	-4.137	-2.445	20.903	14.496	17.504	0.0024	0.0030	0.0029	-0.0061	-0.0058	-0.0058
	(2.638)	(2.758)	(2.782)	11.647	(12.365)	(12.699)	0.0014	(0.0014)	(0.0014)	0.0057	(0.0061)	(0.0062)
2011	5.596	-4.130	-2.035	7.572	1.141	1.875	0.0010	0.0022	0.0018	-0.0038	-0.0035	-0.0028
	(2.630)	(2.747)	(2.771)	11.726	(12.411)	(12.742)	0.0013	(0.0014)	(0.0014)	0.0058	(0.0061)	(0.0063)
2012	11.584	0.875	3.005	9.143	-3.218	-6.412	-0.0008	0.0003	-0.0000	0.0025	0.0055	0.0072
	(2.634)	(2.752)	(2.776)	11.838	(12.570)	(12.972)	0.00135	(0.0014)	(0.0014)	0.0058	(0.0062)	(0.0064)
# Municipalities	5565	5565	5565	323	323	323	5565	5565	5565	323	323	323
Constant	3327	3245	3258	3037	3311	3183	0.116	0.0763	0.0743	0.178	0.0729	0.1152
F statistic	5.17	24.59	36.4	2.21	2.91	2.55	0.96	3.73	3.56	1.53	2.64	1.72
p-value	0.000	0.000	0.000	0.050	0.0123	0.026	0.438	0.002	0.003	0.177	0.022	0.126
Covariates x Linear time trend		X	X		X	X		X	X		X	X
State x Linear time trend			X			X			X			X
N	20,084,702	20,084,702	20,084,702	415,872	415,872	415,872	20,084,702	20,084,702	20,084,702	415,872	415,872	415,872

Notes: Data comes from Brazilian vital statistics for years 2006 to 2013 and from the Ministry of Finance. Each row reports the coefficient on the interaction of an indicator for municipalities that ever received PMM doctors and the labeled year fixed effect. The reported F statistic and corresponding p-value are from a test of the joint significance of the difference between the interaction term and the year fixed effect for all years. Specifications in columns (1), (4), (7), and (10) include a set of

Table 3.5: Sociodemographic and Geographic Predictors of PMM doctors

	PMM doctors per 10,000 residents								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log population density	-0.190	-0.183	-0.166	-0.156	-0.156	-0.153	-0.152	-0.160	-0.149
	(.010)	(.011)	(.011)	(.012)	(.012)	(.013)	(.013)	(.013)	(.014)
Black share of population		-0.013	-0.017	-0.020	-0.020	-0.020	-0.020	-0.020	**-0.008**
		(.003)	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)
Literacy rate			-0.016	-0.003	-0.003	-0.002	-0.002	-0.001	-0.005
			(.002)	(.004)	(.004)	(.004)	(.004)	(.004)	(.005)
MHDI Income				-1.672	-1.736	-1.601	-1.692	-1.796	**-3.224**
				(.437)	(.478)	(.525)	(.530)	(.536)	(.558)
MHDI Longevity					0.227	0.227	0.233	0.218	0.695
					(.702)	(.702)	(.702)	(.700)	(.731)
MHDI Education						-0.257	-0.268	-0.291	0.407
						(.363)	(.363)	(.363)	(.377)
Primary care doctors, 2013							0.001	0.001	**0.001**
							(.000)	(.000)	(.000)
Rural region								-0.131	-0.084
								(.079)	(.079)
North									0.036
									(.073)
Northeast									0.101
									(.066)
Midwest									**0.279**
									(.076)
South									**0.598**
									(.052)
Constant	2.078	2.141	3.446	3.428	3.324	3.291	3.304	3.348	3.571
R^2	0.05	0.05	0.06	0.06	0.06	0.07	0.07	0.07	0.09
N	5,565	5,562	5,562	5,562	5,562	5,562	5,562	5,562	5,562

Data: IBGE and CNES. All specifications include baseline municipal population as measured in 2013. Standard errors clustered at the municipality level are presented in parentheses.

Table 3.6: Test of parallel pre-trends for Continuous Treatment DiD model, Full Sample

	2007	2008	2009	2010	2011	2012	Municip.	Constant	F stat	p-value	N
Full Sample											
					Panel A. Low Intensity vs Control						
Birth weight (grams)	-1.731	-3.695	1.361	-2.979	-1.897	2.957	3,628	3223	22.2	0.000	17,916,978
	(2.828)	(2.824)	(2.842)	(2.849)	(2.837)	(2.840)					
LBW	0.0001	-0.0003	0.0012	0.0030	0.0021	-0.0001	3,628	0.086	2.56	0.025	17,916,978
	(.0020)	(.0015)	(.0015)	(.0015)	(.0015)	(.0015)					
					Panel B. High Intensity vs Control						
Birth weight (grams)	3.456	6.326	3.390	0.991	-1.452	4.577	2,862	3294	7.72	0.000	2,794,884
	(3.389)	(3.382)	(3.411)	(3.413)	(3.407)	(3.417)					
LBW	-0.0030	-0.0041	-0.0005	0.0021	0.0007	-0.0011	2,862	0.064	2.55	0.026	2,794,884
	(.0020)	(.0017)	(.0017)	(.0017)	(.0017)	(.0017)					
					Panel C. Low Intensity vs High Intensity						
Birth weight (grams)	1.530	5.084	0.767	-0.272	-1.683	-2.208	4,640	3257	34.40	0.000	19,457,542
	(1.729)	(1.725)	(1.739)	(1.754)	(1.747)	(1.751)					
LBW	-0.0010	-0.0010	-0.0005	0.0006	0.0005	0.0008	4,640	0.076	3.96	0.001	19,457,542
	(.0010)	(.0001)	(.0009)	(.0009)	(.0009)	(.0009)					

Notes: Data comes from Brazilian vital statistics for years 2006 to 2013 and from the Ministry of Finance. Each column reports the coefficient on the interaction of an indicator for the intensity of PMM treatment and a year fixed effect. The reported F statistic and corresponding p-value are from a test of the joint significance of the difference between the interaction term and the year fixed effect for all years. Intensity of treatment is defined using the mean PMM doctors per capita at the peak of the program, January 2016, such that municipalities who on average received less than the mean were low-treated and those at or above the mean were high-treated. All specifications include a set of baseline municipality controls measured in 2013 that include log gdp, log municipal population, log share of total municipality spending on welfare, share of total municipal spending of municipal gdp, share of municipal residents that receive Bolsa Familia and average Bolsa Familia transfer per recipient interacted with a linear time trend. Incidence of low birth weight is defined as births with birth weight under 2,500 grams, per World Health Organization guidelines. Standard errors are clustered at the municipality level and are included in the parentheses.

Table 3.7: Test of parallel pre-trends for Continuous Treatment DiD model, Rural Sample

	2007	2008	2009	2010	2011	2012	Municip.	Constant	F stat	p-value	N
Rural Sample											
				Low Intensity vs Control							
Birth weight (grams)	-26.496	9.205	-19.202	2.176	-6.947	-23.818	161	3211	1.60	0.156	239,477
	(15.813)	(15.768)	(15.838)	(16.011)	(15.943)	(16.286)					
LBW	0.0178	-0.0049	0.0034	0.0026	0.0090	0.0196	161	0.115	2.73	0.018	239,477
	(.0077)	(.0077)	(.0077)	(.0078)	(.0078)	(.0080)					
				High Intensity vs Control							
Birth weight (grams)	4.090	36.093	3.671	15.968	-6.178	-0.989	213	3280	1.20	0.305	207,719
	(13.461)	(13.478)	(13.581)	(13.728)	(13.834)	(14.117)					
LBW	-0.0028	-0.0224	-0.0054	-0.0054	-0.0051	0.0035	213	0.085	0.65	0.664	207,719
	(.0067)	(.0067)	(.0067)	(.0068)	(.0069)	(.0070)					
				Low Intensity vs High Intensity							
Birth weight (grams)	4.213	1.495	-0.579	-4.922	-5.791	4.334	272	3149	1.62	0.150	384,548
	(7.493)	(7.482)	(7.539)	(7.709)	(7.505)	(7.606)					
LBW	-0.0041	-0.0007	0.0010	0.0016	-0.0007	-0.0073	272	0.104	1.60	0.157	384,548
	(.0037)	(.0037)	(0.0037)	(0.0038)	(0.0037)	(0.0037)					

Notes: Data comes from Brazilian vital statistics for years 2006 to 2013 and from the Ministry of Finance. Each column reports the coefficient on the interaction of an indicator for the intensity of PMM treatment and a year fixed effect. The reported F statistic and corresponding p-value are from a test of the joint significance of the difference between the interaction term and the year fixed effect for all years. Intensity of treatment is defined using the mean PMM doctors per capita at the peak of the program, January 2016, such that municipalities who on average received less than the mean were low-treated and those at or above the mean were high-treated. All specifications include a set of baseline municipality controls measured in 2013 that include log gdp, log municipal population, log share of total municipality spending on welfare, share of total municipal spending of municipal gdp, share of municipal residents that receive Bolsa Familia and average Bolsa Familia transfer per recipient interacted with a linear time trend. Incidence of low birth weight is defined as births with birth weight under 2,500 grams, per World Health Organization guidelines. The rural sample uses the definition from IBGE of rural remote municipalities. This definition of rural takes into account population size and urbanization rate. Standard errors are clustered at the municipality level and are included in the parentheses.

Table 3.8: Effects of PMM on birth weight, All Specifications

	Full Sample					
	OLS (1)	DiD - Binary (2)	DiD - VTT (3)	DiD Cont. (4)	IV (5)	LIML (6)
Birth weight (grams)						
PMM	0.476	-0.425	2.663	0.088	44.180	44.534
	(0.901)	(2.816)	(1.117)	(0.646)	(22.632)	(22.820)
Incidence of low birth weight (¡2,500g)						
PMM	-0.0001	0.0004	-0.0008	-0.0001	-0.0070	-0.0070
	(0.0003)	(0.0010)	(0.0005)	(0.0002)	(0.0053)	(0.0053)
First stage F statistic					33.43	33.43
Baseline covariates x Linear time trend	X	X	X	X	X	X
State x Linear time trend	X	X	X	X	X	X
N	12,288,218	13,847,568	13,847,568	13,847,568	15,367,087	12,291,855

Notes: Data comes from Brazilian vital statistics for years 2006 to 2017. Birth weight is measured in grams. Incidence of low birth weight is defined as births with birth weight under 2,500 grams, per World Health Organization guidelines. The set of controls includes, log share of municipal spending on welfare, and log share of municipal spending as a function of municipal GDP. Columns (1) and (4) define treatment as the number of PMM physicians per capita in a municipality-month. Column (2) defines treated municipalities as those that ever received PMM physicians and thereafter. Column (3) allows for variation in treatment timing such that municipalities are defined as treated only on those months that they have non-zero PMM physicians. Columns (5) and (6) instrument for PMM physicians per capita using the PMM profiles defined at baseline in 2013. Columns (3) and (4) use a 50 percent random sample for estimation. Standard errors clustered at the municipality level are included in the parentheses.

Table 3.9: Effects of PMM on birth weight, All Specifications

	Rural Sample					
	OLS (1)	DiD - Binary (2)	DiD - VTT (3)	DiD Cont. (4)	IV (5)	LIML (6)
Birth weight (grams)						
PMM	3.104	4.202	-3.565	0.435	103.586	103.915
	(2.327)	(9.754)	(4.972)	(2.039)	(56.329)	(56.540)
Incidence of low birth weight (¡2,500g)						
PMM	-0.0006	-0.0008	0.0010	0.0004	-0.0124	-0.0124
	(0.0009)	(0.0037)	(0.0021)	(0.0008)	(0.0126)	(0.0126)
First stage F statistic					8.32	8.32
Baseline covariates x Linear time trend	X	X	X	X	X	X
State x Linear time trend	X	X	X	X	X	X
N	716,533	716,533	716,533	716,533	716,299	716,299

Notes: Data comes from Brazilian vital statistics for years 2006 to 2017. Birth weight is measured in grams. Incidence of low birth weight is defined as births with birth weight under 2,500 grams, per World Health Organization guidelines. The set of controls includes, log share of municipal spending on welfare, and log share of municipal spending as a function of municipal GDP. Columns (1) and (4) define treatment as the number of PMM physicians per capita in a municipality-month. Column (2) defines treated municipalities as those that ever received PMM physicians and thereafter. Column (3) allows for variation in treatment timing such that municipalities are defined as treated only on those months that they have non-zero PMM physicians. Columns (5) and (6) instrument for PMM physicians per capita using the PMM profiles defined at baseline in 2013. Standard errors clustered at the municipality level are included in the parentheses.

Table 3.10: Effects of PMM on birth weight, DiD: Binary

	Full Sample				Rural Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth weight (grams)								
PMM	7.050	4.374	-2.709	-0.425	10.24	5.384	4.63	4.202
	(2.654)	(2.521)	(2.491)	(2.816)	(9.367)	(8.711)	(9.144)	(9.754)
Incidence of low birth weight (¡2,500g)								
PMM	-0.0002	0.0001	0.0007	0.0004	0.000	0.001	0.0004	-0.0008
	(0.001)	(0.0008)	(0.0008)	(0.0010)	(0.004)	(0.004)	(0.0036)	(0.0037)
Yearly Muni. Covariates		X				X		
Baseline covariates x Linear time trend			X	X			X	X
State x Linear time trend				X				X
N	34,628,420	34,628,420	34,628,245	13,847,568	716,533	716,533	716,533	716,533

Notes: Data comes from Brazilian vital statistics for years 2006 to 2017. Birth weight is measured in grams. Incidence of low birth weight is defined as births with birth weight under 2,500 grams, per World Health Organization guidelines. The set of yearly municipality controls includes log gdp, log municipal population, log share of total municipality spending on welfare, share of total municipal spending of municipal gdp, share of municipal residents that receive Bolsa Familia and average Bolsa Familia transfer per recipient. Baseline covariates are those listed above, but measured only in 2013. The most saturated models also include state-specific linear time trends. Note, column (4) uses a 50 percent random sample for estimation as using the full sample was not possible given the computational intensity of the additional fixed effects. Standard errors clustered at the municipality level are included in the parentheses.

Table 3.11: Effects of PMM on birth weight, DiD: Continuous

	Full Sample				Rural Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth weight (grams)								
Physicians per capita	-5.822	-4.437	-0.995	0.088	0.838	-0.379	0.000	0.435
	(0.699)	(0.576)	(0.605)	(0.646)	(1.849)	(1.791)	(1.740)	(2.039)
Incidence of low birth weight (¡2,500g)								
Physicians per capita	0.0005	0.0004	0.0003	-0.0001	0.0001	0.0004	0.0005	0.0004
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0007)	(0.0007)	(0.0007)	(0.0008)
Yearly Muni. Covariates		X				X		
Baseline covariates x Linear time trend			X	X			X	X
State x Linear time trend				X				X
N	34,628,420	34,628,420	34,628,420	13,847,568	716,533	716,533	716,533	716,533

Notes: Treatment is defined as the number of PMM physicians per capita in a given municipality-month. Data comes from Brazilian vital statistics for years 2006 to 2017. Birth weight is measured in grams. Incidence of low birth weight is defined as births with birth weight under 2,500 grams, per World Health Organization guidelines. The set of yearly municipality controls includes log gdp, log municipal population, log share of total municipality spending on welfare, share of total municipal spending of municipal gdp, share of municipal residents that receive Bolsa Familia and average Bolsa Familia transfer per recipient. Baseline covariates are those listed above, but measured only in 2013. The most saturated models also include state-specific linear time trends. Note, column (4) uses a 50 percent random sample for estimation as using the full sample was not possible given the computational intensity of the additional fixed effects. Standard errors clustered at the municipality level are included in the parentheses.

Table 3.12: Effects of PMM on birth weight, OLS

	Full Sample				Rural Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth weight (grams)								
Physicians per capita	11.320	2.253	3.671	0.476	-6.597	0.458	-0.360	3.104
	(2.129)	(0.940)	(1.170)	(0.901)	(1.883)	(2.215)	(2.919)	(2.327)
Incidence of low birth weight (¡2,500g)								
Physicians per capita	-0.0026	-0.0005	-0.0004	-0.0001	0.0006	-0.0008	0.0001	-0.0006
	(0.0005)	(0.0002)	(0.0003)	(0.0003)	(0.0005)	(0.0008)	(0.0009)	(0.0009)
Yearly Muni. Covariates		X				X		
Baseline covariates x Linear time trend			X	X			X	X
State x Linear time trend				X				X
N	15,367,087	15,367,087	15,367,087	12,288,218	716,533	716,533	716,533	716,533

Notes: Treatment is defined as the number of PMM physicians per capita in a given municipality-month. Data comes from Brazilian vital statistics for years 2006 to 2017. Birth weight is measured in grams. Incidence of low birth weight is defined as births with birth weight under 2,500 grams, per World Health Organization guidelines. The set of yearly municipality controls includes log share of total municipality spending on welfare and log share of total municipal spending of municipal gdp. Baseline covariates are those listed above, but measured only in 2013. The most saturated models also include state-specific linear time trends. Estimates in columns (1) - (4) use a 50 percent random sample. Standard errors clustered at the municipality level are included in the parentheses.

Table 3.13: PMM Profiles and Physicians per capita

	Full Sample				Rural Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Profile 3	-0.0877	0.0568	0.1010	0.0793				
	(0.0217)	(0.0185)	(0.0169)	(0.0157)				
Profile 5	0.0174	0.1394	0.1254	0.1184				
	(0.0196)	(0.0264)	(0.0213)	(0.0176)				
Profile 6	0.0952	0.1108	0.1206	0.1098	0.1935	0.2174	0.2901	0.1751
	(0.0177)	(0.0238)	(0.0195)	(0.0173)	(0.0762)	(0.1229)	(0.0989)	(0.0817)
Profile 7	0.2225	0.1242	0.2090	0.1906	0.2270	0.1617	0.3255	0.2319
	(0.0113)	(0.0239)	(0.0200)	(0.0168)	(0.0474)	(0.0872)	(0.0850)	(0.0598)
F statistic	113.83	10.35	34.11	33.43	11.78	2.13	7.86	8.32
Yearly Muni. Covariates		X				X		
Baseline covariates x Linear time trend			X	X			X	X
State x Linear time trend				X				X
N	15,367,087	15,367,087	15,367,087	15,367,087	716,299	716,299	716,299	716,299

Notes: Data comes from Brazilian vital statistics for years 2006 to 2017.The set of yearly municipality controls includes log gdp, log municipal population, log share of total municipality spending on welfare, share of total municipal spending of municipal gdp, share of municipal residents that receive Bolsa Familia and average Bolsa Familia transfer per recipient. Baseline covariates are those listed above, but measured only in 2013. The most saturated models also include state-specific linear time trends. All specifications show results of instrumenting the municipal number of PMM physicians per capita with PMM profiles. Municipalities that were classified using their PAB (federal health services funding) are excluded from the estimation. Standard errors clustered at the municipality level are included in the parentheses. Note, full sample estimates are from a 50 percent random sample for estimation as using the full sample was not possible given

Table 3.14: Effects of PMM on birth weight, DiD: Binary with variation in treatment timing

	Full Sample				Rural Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth weight (grams)								
PMM	4.911	3.462	0.491	2.663	0.675	-3.723	-4.649	-3.565
	(1.142)	(1.027)	(0.981)	(1.117)	(4.928)	(4.712)	(4.734)	(4.972)
Incidence of low birth weight (¡2,500g)								
PMM	-0.0008	-0.0007	-0.0004	-0.0008	0.0014	0.0024	0.0023	0.0010
	(0.0003)	(0.0003)	(0.0003)	(0.0005)	(0.0020)	(0.0020)	(0.0020)	(0.0021)
Yearly Muni. Covariates		X				X		
Baseline covariates x Linear time trend			X	X			X	X
State x Linear time trend				X				X
N	34,628,420	34,628,420	34,628,420	13,847,568	716,533	716,533	716,533	716,533

Notes: Treatment ”turns on” for an observation if the municipality of birth has PMM physicians present that month, and off otherwise. Data comes from Brazilian vital statistics for years 2006 to 2017. Birth weight is measured in grams. Incidence of low birth weight is defined as births with birth weight under 2,500 grams, per World Health Organization guidelines. The set of yearly municipality controls includes log gdp, log municipal population, log share of total municipality spending on welfare, share of total municipal spending of municipal gdp, share of municipal residents that receive Bolsa Familia and average Bolsa Familia transfer per recipient. Baseline covariates are those listed above, but measured only in 2013. The most saturated models also include state-specific linear time trends. Note, column (4) uses a 50 percent random sample for estimation as using the full sample was not possible given the computational intensity of the additional fixed effects. Standard errors are clustered at the municipality level are included in the parentheses.

Table 3.15: Effects of PMM on birth weight, All Specifications

	Rural Sample					
	OLS (1)	DiD - Binary (2)	DiD - VTT (3)	DiD Cont. (4)	IV (5)	LIML (6)
Birth weight (grams)						
PMM	3.104	4.202	-3.565	0.435	103.586	103.915
	(2.327)	(9.754)	(4.972)	(2.039)	(56.329)	(56.540)
Incidence of low birth weight (¡2,500g)						
PMM	-0.0006	-0.0008	0.0010	0.0004	-0.0124	-0.0124
	(0.0009)	(0.0037)	(0.0021)	(0.0008)	(0.0126)	(0.0126)
First stage F statistic					8.32	8.32
Baseline covariates x Linear time trend	X	X	X	X	X	X
State x Linear time trend	X	X	X	X	X	X
N	716,533	716,533	716,533	716,533	716,299	716,299

Notes: Data comes from Brazilian vital statistics for years 2006 to 2017. Birth weight is measured in grams. Incidence of low birth weight is defined as births with birth weight under 2,500 grams, per World Health Organization guidelines. The set of controls includes, log share of municipal spending on welfare, and log share of municipal spending as a function of municipal GDP. Columns (1) and (4) define treatment as the number of PMM physicians per capita in a municipality-month. Column (2) defines treated municipalities as those that ever received PMM physicians and thereafter. Column (3) allows for variation in treatment timing such that municipalities are defined as treated only on those months that they have non-zero PMM physicians. Columns (5) and (6) instrument for PMM physicians per capita using the PMM profiles defined at baseline in 2013. Standard errors clustered at the municipality level are included in the parentheses.

Table 3.16: Effects of PMM on birth weight, Instrumental Variables

	Full Sample				Rural Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth weight (grams)								
Physicians per capita	228.502	65.230	95.668	44.180	-20.912	-65.959	13.940	103.586
	(16.085)	(37.565)	(24.569)	(22.632)	(38.830)	(71.368)	(40.772)	(56.329)
Incidence of low birth weight (¡2,500g)								
Physicians per capita	-0.0503	-0.0138	-0.0139	-0.0070	0.0360	0.0435	0.0068	-0.0124
	(0.0034)	(0.0090)	(0.0056)	(0.0053)	(0.0113)	(0.0302)	(0.0094)	(0.0126)
First stage F statistic	113.83	10.35	34.11	33.43	11.78	2.13	7.86	8.32
Yearly Muni. Covariates		X				X		
Baseline covariates x Linear time trend			X	X			X	X
State x Linear time trend				X				X
N	15,367,087	15,367,087	15,367,087	15,367,087	716,299	716,299	716,299	716,299

Notes: Data comes from Brazilian vital statistics for years 2006 to 2017. Birth weight is measured in grams. Incidence of low birth weight is defined as births with birth weight under 2,500 grams, per World Health Organization guidelines. The set of yearly municipality controls includes the log share of total municipality spending on welfare and the log share of total municipal spending of municipal gdp. Note, columns (1) - (4) use a 50 percent random sample for estimation. All specifications show results of instrumenting the municipal number of PMM physicians per capita with PMM profiles. Municipalities that were classified using their PAB (federal health services funding) are excluded from the estimation. Standard errors clustered at the municipality level are included in the parentheses.

Table 3.17: Effects of PMM on birth weight

(a) Limited Information Maximum Likelihood (LIML)

	Full Sample				Rural Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth weight (grams)								
Physicians per capita	229.354 (16.183)	67.932 (39.235)	96.453 (24.793)	44.534 (22.820)	-21.093 (39.316)	-67.674 (73.638)	14.144 (41.349)	103.915 (56.540)
Incidence of low birth weight (¡2,500g)	-0.0503 (0.0035)	-0.0138 (0.0091)	-0.0139 (0.0056)	-0.0070 (0.0053)	0.0360 (0.0113)	0.0435 (0.0302)	0.0068 (0.0094)	-0.0124 (0.0126)
Physicians per capita								
First stage F statistic	113.83	10.35	34.11	33.43	11.78	2.13	7.86	8.32
Yearly Muni. Covariates		X				X		
Baseline covariates x Linear time trend			X	X			X	X
State x Linear time trend				X				X
N	15,366,255	15,366,255	15,366,255	12,291,855	716,299	716,299	716,299	716,299

Notes: Data comes from Brazilian vital statistics for years 2006 to 2017. Birth weight is measured in grams. Incidence of low birth weight is defined as births with birth weight under 2,500 grams, per World Health Organization guidelines. The set of yearly municipality controls includes the log share of total municipality spending on welfare and the log share of total municipal spending of municipal gdp. Note, columns (1) - (4) use a 50 percent random sample for estimation. All specifications show results of instrumenting the municipal number of PMM physicians per capita with PMM profiles. Municipalities that were classified using their PAB (federal health services funding) are excluded from the estimation. Standard errors clustered at the municipality level are included in the parentheses.

Appendix A: Data Appendix: Increasing Supply of Physicians and Birth Outcomes

Data Appendix

SINASC data change

The change to the SINASC data included changes in the composition of some variables and introduction of new variables. All of the changes are reported by [140]. Existing variables that underwent a change in collection method were mother's age at birth, mother's educational attainment, number of past children born alive, number of past children born deceased, duration of current pregnancy, and the number prenatal consultations.

Mother's age

Previously, mother's age was recorded from a direct question to the mother about her age. The new version also asks mother's date of birth. Then mother's age is calculated as the difference between her date of birth and the date of birth of the child. However, when mother's date of birth is unknown, this field is completed using the old method. In 2011, 99% of the mother's age fields are completed using the date of birth. Ultimately, there exist minimal fluctuations in the composition of the age of mothers from 2011 and the previous years.

Mother's education attainment

The previous method collected mother's educational attainment through a categorical variable ranging from none, 1-3 years, 4-7 years, 8-11 years, and 12 or more years of education. The new method also reports the level of schooling completed using the original variable to better match the surveys conducted by the IBGE. From 2010 to 2011, there is a slight decrease in those reporting 12 years or more with a proportional increase in those reporting 8 to 11 years.

Marital Status

The new Brazilian civil code began to include a ”stable union” as a recognized civil status, albeit not a ”marital status”. Consequently, this value was added to the options for marital status in SINASC. With this new option, share of mothers reporting themselves as single falls by the proportion that stable union rises, about 14 percent.

Race

New model includes a question about the mother's race, not just the newborn's race as in the past. From this change, there is a decrease in percent white with proportional increase in percent black, while brown remains at its 2020 level. These values are very close to those reported by IBGE.

Previous children

While the old survey asked only number live and deceased births, the new survey asks more information about past pregnancies. New questions include number of past pregnancies, number of vaginal births, number of c-sections, and number of miscarriages/abortions. These new variables are reported starting in 2012. The report by MoH detects a decrease in no previous deceased births matching an increase in 1 previous deceased birth of about 5 percent.

Gestational Length

Prior to 2010, gestational length was captured by a categorical variable with bins of about 4 weeks. Since then, gestational length is calculated using the new question about the mother's date of last menstrual cycle (LMP) and reported as number of weeks, a continuous variable. Weeks of gestation is calculated using the LMP and the date of birth of newborn and it is recorded both in the continuous and categorical gestation variables. When LMP is unknown, the form asks the mother directly the number of weeks of gestation. In 2011, gestation is recorded for 96.1% of births 2011conso. Resulting from the change, missing values increased by almost 4%, and premature births increased throughout the country. The MoH argues this increase is not due to a change in actual premature births, but from increased precision in the estimation of gestational length. This is supported by the fact that other national surveys indicated higher levels of premature births than

SINASC.

Prenatal visits

Similar to gestational length, prenatal visits was collected in a categorical variable that changed to a continuous precise number of visits value. The MoH reports no change to the historical trends in prenatal consultations from this change.

Bibliography

[1] Gary S Becker. Crime and punishment: An economic approach. *Journal of political economy*, 76(2):169–217, 1968.

[2] Stores score on pension payday. *Mail and Guardian*, 3 Feb 2012. Available at: https://mg.co.za/article/2012-02-03-stores-score-on-pension-payday/(Accessed: July 20, 2023).

[3] Fernando Borraz, Ignacio Munyo, et al. Conditional cash transfers and crime: Higher income but also better loot. *Economics Bulletin*, 40(2):1804–1813, 2020.

[4] Sonja Raasch. Sassa acting ceo concerned following cala pay point robbery. *The Rep*, 11 Jan 2018. Available at: https://www.therep.co.za/2018/01/11/sassa-acting-ceo-concerned-following-cala-pay-point-robbery/ (Accessed: July 10, 2023).

[5] Nwabisa Msutwana-Stemela. South africa: Cops tighten security at pay points. *allAfrica*, 11 Jan 2009. Available at: https://allafrica.com/stories/200901121075.html (Accessed: July 10, 2023).

[6] Cebelihle Mthethwa. Plato: Sassa needs security upgrade. *News24*, 1 Jun 2014. Available at: https://www.news24.com/News24/Plato-Sassa-needs-security-upgrade-20140601 (Accessed: July 10, 2023).

[7] Suthentira Govender. Elderly kzn woman killed for her pension by 'young man who knew her'. *Sowetan Live*, 12 May 2020. Available at: https://www.sowetanlive.co.za/news/south-africa/2020-05-12-elderly-kzn-woman-killed-for-her-pension-by-young-man-who-knew-her (Accessed: July 10, 2023).

[8] Nation Nyoka. Woman gets 20 years in jail for killing husband over pension money. *News24*, 4 Oct 2020. Available at: https://www.news24.com/News24/woman-gets-20-years-in-jail-for-killing-husband-over-pension-money-20171004 (Accessed: July 10, 2023).

[9] Pensioner killed as robbers hit pension point in diepkloof. *ENCA*, 3 May 2016. Available at: https://www.enca.com/south-africa/pensioner-killed-as-robbers-hit-pension-point-in-diepkloof (Accessed: July 19, 2023).

[10] Finmark Trust. Sassa grant distribution: Improving the financial capability of grant recipients. Technical report, 2018. Available at: https://finmark.org.za/system/documents/files/000/000/276/original/SASSA_Grant_Recipients_-_Improving_the_Financial_Capability.pdf?1605614633 (Accessed: May 20, 2023).

[11] Ken Miyajima. The link between social grants and employment in south africa. *Selected Issues Papers*, 2023(039), 2023.

[12] Reza Varjavand. Growing underground economy; the evidence, measures, and the consequences. *Journal of International Management Studies*, 11(3):133–142, 2011.

[13] Neil Wallace. The case for imposing cashlessness: A review article. *Journal of Economic Literature*, 56(4):1587–91, 2018.

[14] K Rogoff. The curse of cash: How large-denomination bills aid crime and tax evasion and constrain monetary policy, 2017.

[15] David R Warwick. Reducing crime by eliminating cash. 1993.

[16] William Alex Pridemore, Sean Patrick Roche, and Meghan L Rogers. Cashlessness and street crime: A cross-national study of direct deposit payment and robbery rates. *Justice Quarterly*, 35(5):919–939, 2018.

[17] RT Naylor. A ruse by any other name: The underground economy. *Challenge*, 48(6):32–49, 2005.

[18] Friedrich Schneider and Dominik H Enste. *The shadow economy: An international survey*. Cambridge University Press, 2013.

[19] Erdal Tekin, Volkan Topalli, Chandler McClellan, and Richard Wright. Liquidating crime with illiquidity: How switching from cash to credit can stop street crime. *CESifo DICE Report*, 12(3):45–50, 2014.

[20] Nestor Gandelman, Ignacio Munyo, and Emanuel Schertz. Cash and crime. In *Mimeo*. Universidad ORT Uruguay, 2019.

[21] Richard Wright, Erdal Tekin, Volkan Topalli, Chandler McClellan, Timothy Dickinson, and Richard Rosenfeld. Less cash, less crime: Evidence from the electronic benefit transfer program. *The Journal of Law and Economics*, 60(2):361–383, 2017.

[22] Laura E Armey, Jonathan Lipow, and Natalie J Webb. The impact of electronic financial payments on crime. *Information Economics and Policy*, 29:46–57, 2014.

[23] Jillian B Carr and Analisa Packham. Snap benefits and crime: Evidence from changing disbursement schedules. *Review of Economics and Statistics*, 101(2):310–325, 2019.

[24] Neal Shover. *Great pretenders: Pursuits and careers of persistent thieves*. Routledge, 2018.

[25] Richard T Wright and Scott H Decker. Burglars on the job, streetlife and residential break-ins. *International Journal of the Sociology of Law*, 23(3):299, 1995.

[26] Richard T Wright and Scott H Decker. *Armed robbers in action: Stickups and street culture*. UPNE, 1997.

[27] Sara Markowitz. Alcohol, drugs and violent crime. *International Review of Law and economics*, 25(1):20–44, 2005.

[28] Kathleen Auerhahn and RN Parker. Alcohol, drugs, and violence. *Annual Review of Sociology*, 24(1):291–31, 1998.

[29] Dominic J Parrott and Christopher I Eckhardt. Effects of alcohol on human aggression. *Current opinion in psychology*, 19:1–5, 2018.

[30] Ted R Miller, David T Levy, Mark A Cohen, and Kenya LC Cox. Costs of alcohol and drug-involved crime. *Prevention Science*, 7:333–342, 2006.

[31] Jacky M Jennings, Adam J Milam, Amelia Greiner, C Debra M Furr-Holden, Frank C Curriero, and Rachel J Thornton. Neighborhood alcohol outlets and the association with violent crime in one mid-atlantic city: the implications for zoning policy. *Journal of Urban Health*, 91:62–71, 2014.

[32] Brett Watson, Mouhcine Guettabi, and Matthew Reimer. Universal cash and crime. *Review of Economics and Statistics*, 102(4):678–689, 2020.

[33] Carlos Dobkin and Steven L Puller. The effects of government transfers on monthly cycles in drug abuse, hospitalization and mortality. *Journal of Public Economics*, 91(11-12):2137–2157, 2007.

[34] Elena Castellari, Chad Cotti, John Gordanier, and Orgul Ozturk. Does the timing of food stamp distribution matter? a panel-data analysis of monthly purchasing patterns of us households. *Health economics*, 26(11):1380–1393, 2017.

[35] Cody Tuttle. Snapping back: Food stamp bans and criminal recidivism. *American Economic Journal: Economic Policy*, 11(2):301–327, 2019.

[36] Mkhuseli Sizani and Daniel Steyn. Closure of sassa pay points 'catastrophic' for pensioners in rural areas. *News24*, 15 Sep 2022. Available at: https://www.news24.com/citypress/news/closure-of-sassa-pay-points-catastrophic-for-pensioners-in-rural-areas-20220915 (Accessed: February 10, 2023).

[37] PYMNTS. Global cash index - south africa edition. Technical report, 2017. Available at: http://pymnts.fetchapp.com/files/d8764a (Accessed: September 25, 2022).

[38] In south africa, cash is consumers' hands-down choice. *PYMNTS*, 30 Jun 2017. Available at: https://www.pymnts.com/cash/2017/south-african-consumers-pick-cash/(Accessed: October 19, 2021).

[39] Lynette Maart, Angie Richardson, and Rachel Bukasa. Hands off our grants: the black sash's long battle for justice. *Ground Up*, 19 May 2022. Available at: https://www.groundup.org.za/article/hands-off-our-grants-the-black-sashs-long-battle-for-justice/ (Accessed: July 10, 2023).

[40] Mubeen Bandeker. Grant deductions are immoral: Black sash. *The Voice of the Cape*, 9 Sept 2014. Available at: https://www.vocfm.co.za/grant-deductions-are-immoral-black-sash/ (Accessed: July 10, 2023).

[41] Spotlight on social grants: Sassa's bid to stop illegal deductions. *Ground Up*, 7 Oct 2015. Available at: https://www.groundup.org.za/article/spotlight-social-grants-sassas-bid-stop-illegal-deductions_3372/ (Accessed: July 10, 2023).

[42] South african bank fees under fire. *Business Tech*, 15 Feb 2015. Available at: https://businesstech.co.za/news/banking/79778/south-african-bank-fees-under-fire/ (Accessed: July 10, 2023).

[43] Adam Ikdal, Euvin Naidoo, Adrien Portafaix, Joshua Hendrickson, Alex Boje, Darryn Rabec, and Klaus Kessler. Improving financial inclusion in south africa. Technical report, 2017. Available at: https://www.bcg.com/publications/2017/globalization-improving-financial-inclusion-south-africa (Accessed: July 19, 2023).

[44] First National Bank. Understand your bank fees. Technical report, 2014. Available at: https://www.fnb.co.za/downloads/pricing-guide/Personal_Pricing.pdf (Accessed: July 10, 2023).

[45] Velani Ludini. Old people and babies spend night outside in queue for sassa services. *GroundUp*, 26 Apr 2019. Available at: https://www.groundup.org.za/article/old-people-and-babies-spend-night-outside-queue-sassa-services/).

[46] HelpAge International. Accountability in social pension programmes: A baseline mapping of the old age grant in south africa. Technical report, 2014. Available at: https://www.helpage.org/silo/files/accountability-mapping-in-social-pension-programmes-a-baseline-mapping-of-the-old-age-grant-in-south-africa.pdf (Accessed: July 19, 2023).

[47] South African Police Service. Crime situation in republic of south africa twelve (12) months. Technical report, 2020. Available at: https://www.saps.gov.za/services/april_to_march_2019_20_presentation.pdf (Accessed: October 20, 2021).

[48] J van Dijk, J van Kesteren, and Paul Smit. Criminal victimisation in international perspective. 2007.

[49] Mark Shaw, Jan Van Dijk, and Wolfgang Rhomberg. Determining trends in global crime and justice: An overview of results from the united nations surveys of crime trends and operations of criminal justice systems. In *Forum on crime and society*, volume 3, pages 35–63. United Nations Office on Drugs and Crime, 2003.

[50] Michael Rand and Shannan Catalano. Criminal victimization, 2006. *Ncj*, 219413, 2007.

[51] Jennifer L Truman, Lynn Langton, and Michael Planty. Criminal victimization, 2015. *Washington, DC*, 2015.

[52] Richard B Felson, Steven F Messner, Anthony W Hoskin, and Glenn Deane. Reasons for reporting and not reporting domestic violence to the police. *Criminology*, 40(3):617–648, 2002.

[53] Richard R Bennett and R Bruce Wiegand. Observations on crime reporting in a developing nation. *Criminology*, 32(1):135–148, 1994.

[54] Eric P Baumer. Neighborhood disadvantage and police notification by victims of violence. *Criminology*, 40(3):579–616, 2002.

[55] Carmen M Gutierrez and David S Kirk. Silence speaks: The relationship between immigration and the underreporting of crime. *Crime & Delinquency*, 63(8):926–950, 2017.

[56] Herbert B Asher. Some consequences of measurement error in survey data. *American Journal of Political Science*, pages 469–485, 1974.

[57] Donald Edward Joseph. *An assessment of the strengths and weaknesses of the South African Social Security Agency in the Northern and Western Cape Provinces*. PhD thesis, North-West University, 2012.

[58] C Fritz Foley. Welfare payments and crime. *The review of Economics and Statistics*, 93(1):97–112, 2011.

[59] Brian A Jacob and Lars Lefgren. Are idle hands the devil's workshop? incapacitation, concentration, and juvenile crime. *American economic review*, 93(5):1560–1577, 2003.

[60] Lebohang Letsela, Renay Weiner, Mitzy Gafos, and Katherine Fritz. Alcohol availability, marketing, and sexual health risk amongst urban and rural youth in south africa. *AIDS and Behavior*, 23:175–189, 2019.

[61] Late salary payments – know your rights. *Polity*, 27 March 2019. Available at: https://www.polity.org.za/article/late-salary-payments-know-your-rights-2019-03-27#:~:text=However%2C%20the%20majority%20of%20companies,African%20Payroll%20Association%20(SAPA).&text=The%20law%20stipulates%20that%20employers,end%20of%20a%20pay%20period.(Accessed: May 18, 2023).

[62] Lekwa teemane Local Municipality. Pay day policy june 2022. Technical report, 2022. Available at: https://www.lekwateemane.co.za/public/images/uploads/2021/2021Policies/LTLM-Payday%20policy%202021%20done.doc.pdf (Accessed: May 18, 2023).

[63] For many south africans, payday falls after black friday this year - here's why you will still get the deals. *News24*, 17 Nov 2018. Available at: https://www.news24.com/news24/bi-archive/a-late-payday-wont-affect-black-friday-deals-heres-why-2018-11(Accessed: May 18, 2023).

[64] National Center for Science and Engineering Statistics. Doctorate recipients from u.s. universities: 2021. Technical Report NSF 23-300, Alexandria, VA, 2022.

[65] Hilary A Takiff, Diana T Sanchez, and Tracie L Stewart. What's in a name? the status implications of students' terms of address for male and female professors. *Psychology of Women Quarterly*, 25(2):134–144, 2001.

[66] Lillian MacNell, Adam Driscoll, and Andrea N Hunt. What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education*, 40(4):291–303, 2015.

[67] Julia A Files, Anita P Mayer, Marcia G Ko, Patricia Friedrich, Marjorie Jenkins, Michael J Bryan, Suneela Vegunta, Christopher M Wittich, Melissa A Lyle, Ryan Melikian, et al. Speaker introductions at internal medicine grand rounds: forms of address reveal gender bias. *Journal of Women's Health*, 26(5):413–419, 2017.

[68] Natalie Wexler. What's really behind the flap over jill biden's doctorate, 2020.

[69] Tonya Russell. Dr. jill biden deserves her title. saying otherwise demeans teachers and community colleges, 2020.

[70] Joseph Epstein. Is there a doctor in the white house? not if you need an m.d., 2020.

[71] Stav Atir and Melissa J Ferguson. How gender determines the way we speak about professionals. *Proceedings of the National Academy of Sciences*, 115(28):7278–7283, 2018.

[72] Gender Institute for Global Women's Leadership. Gender balance of expert sources in news coverage of covid-19. Technical report, King's College London, 2020.

[73] Miami University Libraries. AP Style Guide, 2022.

[74] E. C. Long., J. Pugel, C. Giray, J. T. Scott, and D. M. Crowley. The role of credibility in communicating substance use research to federal and state policymakers. *Alcoholism: Clinical and Experimental Research*, 45, 2021.

[75] Paula Chatterjee and Rachel M Werner. Gender disparity in citations in high-impact journal articles. *JAMA Network Open*, 4(7):e2114509–e2114509, 2021.

[76] Erin G Teich, Jason Z Kim, Christopher W Lynn, Samantha C Simon, Andrei A Klishin, Karol P Szymula, Pragya Srivastava, Lee C Bassett, Perry Zurn, Jordan D Dworkin, et al. Citation inequity and gendered citation practices in contemporary physics. *Nature Physics*, 18(10):1161–1170, 2022.

[77] Daniel Maliniak, Ryan Powers, and Barbara F Walter. The gender citation gap in international relations. *International Organization*, 67(4):889–922, 2013.

[78] Michelle L Dion, Jane Lawrence Sumner, and Sara McLaughlin Mitchell. Gendered citation patterns across political science and social science methodology fields. *Political analysis*, 26(3):312–327, 2018.

[79] Molly M King, Carl T Bergstrom, Shelley J Correll, Jennifer Jacquet, and Jevin D West. Men set their own cites high: Gender and self-citation across fields and over time. *Socius*, 3:2378023117738903, 2017.

[80] Newly Paul, Mingxiao Sui, and Kathleen Searles. Look who's writing: How gender affects news credibility and perceptions of news relevance. *Journalism & Mass Communication Quarterly*, 99(1):183–212, 2022.

[81] Cory L Armstrong and Melinda J McAdams. Blogs of information: How gender cues and individual motivations influence perceptions of credibility. *Journal of Computer-Mediated Communication*, 14(3):435–456, 2009.

[82] Cory L. Armstrong and Michelle R. Nelson. How newspaper sources trigger gender stereotypes. *Journalism & Mass Communication Quarterly*, 82(4):820–837, 2005.

[83] Elena Klaas and Mark Boukes. A woman's got to write what a woman's got to write: the effect of journalist's gender on the perceived credibility of news articles. *Feminist Media Studies*, 22(3):571–587, 2022.

[84] Alice H. Wu. Gendered language on the economics job market rumors forum. *American Economic Review*, 108:175–79, May 2018.

[85] Anaïs Llorens, Athina Tzovara, Ludovic Bellier, Ilina Bhaya-Grossman, Aurélie Bidet-Caulet, William K. Chang, Zachariah R. Cross, Rosa Dominguez-Faus, Adeen Flinker, Yvonne Fonken, Mark A. Gorenstein, Chris Holdgraf, Colin W. Hoy, Maria V. Ivanova, Richard T. Jimenez, Soyeon Jun, Julia W.Y. Kam, Celeste Kidd, Enitan Marcelle, Deborah Marciano, Stephanie Martin, Nicholas E. Myers, Karita Ojala, Anat Perry, Pedro Pinheiro-Chagas, Stephanie K. Riès, Ignacio Saez, Ivan Skelin, Katarina Slama, Brooke Staveland, Danielle S. Bassett, Elizabeth A. Buffalo, Adrienne L. Fairhall, Nancy J. Kopell, Laura J. Kray, Jack J. Lin, Anna C. Nobre, Dylan Riley, Anne-Kristin Solbakk, Joni D. Wallis, Xiao-Jing Wang, Shlomit Yuval-Greenberg, Sabine Kastner, Robert T. Knight, and Nina F. Dronkers. Gender bias in academia: A lifetime problem that needs solutions. *Neuron*, 109(13):2047–2074, 2021.

[86] Rosemary Morgan, Kate Hawkins, and Jamie Lundine. The foundation and consequences of gender bias in grant peer review processes. *CMAJ*, 190(16):E487–E488, 2018.

[87] Shawn Khan, Abirami Kirubarajan, Tahmina Shamsheri, Adam Clayton, and Geeta Mehta. Gender bias in reference letters for residency and academic medicine: a systematic review. *Postgraduate Medical Journal*, 2021.

[88] Laurie A Rudman. Self-promotion as a risk factor for women: the costs and benefits of counterstereotypical impression management. *Journal of personality and social psychology*, 74(3):629, 1998.

[89] Laurie Rudman and Julie Phelan. Backlash effects for disconfirming gender stereotypes in organizations. *Research in Organizational Behavior - RES ORGAN BEH*, 28:61–79, 12 2008.

[90] Alexandra N. Fisher, Danu Anthony Stinson, and Anastasija Kalajdzic. Unpacking backlash: Individual and contextual moderators of bias against female professors. *Basic and Applied Social Psychology*, 41(5):305–325, 2019.

[91] Laurie A Rudman, Corinne A Moss-Racusin, Julie E Phelan, and Sanne Nauts. Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, 48(1):165–179, 2012.

[92] Jerold Laguilles, Elizabeth Williams, and Daniel Saunders. Can lottery incentives boost web survey response rates? findings from four experiments. *Research in Higher Education*, 52:537–553, 08 2010.

[93] Alauna Safarpour, Sarah Sunn Bush, and Jennifer Hadden. Participation incentives in a survey of international non-profit professionals. *Research & Politics*, 9(3):20531680221125723, 2022.

[94] Anja S Göritz and Susanne C Luthe. Lotteries and study results in market research online panels. *International Journal of Market Research*, 55(5):611–626, 2013.

[95] Anja S Göritz. Incentives in web studies: Methodological issues and a review. *International Journal of Internet Science*, 1(1):58–70, 2006.

[96] Janet K. Swim, Kathryn J. Aikin, Wayne S. Hall, and Barbara A. Hunter. Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology*, 68(2):199–214, February 1995.

[97] James C McCroskey and Thomas J Young. Ethos and credibility: The construct and its measurement after three decades. *Communication Studies*, 32(1):24–34, 1981.

[98] Jacklyn E Nagle, Stanley L Brodsky, and Kaycee Weeter. Gender, smiling, and witness credibility in actual trials. *Behavioral sciences & the law*, 32(2):195–206, 2014.

[99] Luke Lei Zhu, Karl Aquino, and Abhijeet K Vadera. What makes professors appear credible: The effect of demographic characteristics and ideological beliefs. *Journal of Applied Psychology*, 101(6):862, 2016.

[100] Karin Boczek, Leyla Dogruel, and Christiana Schallhorn. Gender byline bias in sports reporting: Examining the visibility and audience perception of male and female journalists in sports coverage. *Journalism*, page 14648849211063312, 2022.

[101] Jeffrey R Kling, Jeffrey B Liebman, and Lawrence F Katz. Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119, 2007.

[102] Dalia Ghanem, Sarojini Hirshleifer, and Karen Ortiz-Becerra. Testing attrition bias in field experiments. 2021.

[103] Joel L Horowitz and Charles F Manski. Nonparametric analysis of randomized experiments with missing covariate and outcome data. *Journal of the American statistical Association*, 95(449):77–84, 2000.

[104] David S Lee. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3):1071–1102, 2009.

[105] Corinne A Moss-Racusin, John F Dovidio, Victoria L Brescoll, Mark J Graham, and Jo Handelsman. Science faculty's subtle gender biases favor male students. *Proceedings of the national academy of sciences*, 109(41):16474–16479, 2012.

[106] Ministério da Saúde. *Programa Mais Médicos – Dois anos: Mais Saúde para os Brasileiros*. Secretaria de Gestão do Trabalho e da Educação na Saúde, 2015.

[107] Ministry of Health. *Programma Mais Medicos*. Secretaria de Gestão do Trabalho e da Educação na Saúde: Departamento de Planejamento e Regulação da Provisão de Profissionais de Saúde, 2017.

[108] John M Kirk, Emily J Kirk, and Chris Walker. Mais médicos: Cuba's medical internationalism programme in brazil. *Bulletin of Latin American Research*, 35(4):467–480, 2016.

[109] El Nuevo Herald. Brasil defiende la contratación de médicos cubanos . 2013.

[110] Mônica Sampaio de Carvalho and Maria Fátima de Sousa. Como o brasil tem enfrentado o tema provimento de médicos? *Interface-Comunicação, Saúde, Educação*, 17(47):913–926, 2013.

[111] P Kiernan. Brazil to Import 4000 Cuban Doctors to Work in Needy Areas'. *Wall Street Journal*, 2013.

[112] Anthony Boadle. Cuban doctors tend to brazil's poor, giving rousseff a boost. *Reuters, available from: http://www. reuters. com/article/2013/12/01/us-brazil-doctors-cuba-idUSBRE9B005720131201 [accessed 7 October 2014]*, 2013.

[113] Yara Aquino. Em protesto contra Mais Médicos, profissionais paralisam atividades em vários estados. *Agencia Brasil*.

[114] Letícia Faria de Carvalho Nunes. *Essays in health economics*. PhD thesis, 2019.

[115] Sábado Nicolau Girardi, Ana Cristina de Sousa van Stralen, Joana Natalia Cella, Lucas Wan Der Maas, Cristiana Leite Carvalho, and Erick de Oliveira Faria. Impact of the mais médicos (more doctors) program in reducing physician shortage in brazilian primary healthcare. *Ciencia & saude coletiva*, 21:2675–2684, 2016.

[116] James Macinko, Barbara Starfield, and Leiyu Shi. Quantifying the health benefits of primary care physician supply in the united states. *International journal of health services*, 37(1):111–126, 2007.

[117] Jostein Grytten and Rune Sørensen. Type of contract and supplier-induced demand for primary physicians in norway. *Journal of health economics*, 20(3):379–393, 2001.

[118] Toshiaki Iizuka and Yasutora Watanabe. The impact of physician supply on the healthcare system: Evidence from japan's new residency program. *Health economics*, 25(11):1433–1447, 2016.

[119] Miranda Laurant, Mieke van der Biezen, Nancy Wijers, Kanokwaroon Watananirun, Evangelos Kontopantelis, and Anneke JAH van Vught. Nurses as substitutes for doctors in primary care. *Cochrane Database of Systematic Reviews*, (7), 2018.

[120] Niteesh K Choudhry, Robert H Fletcher, and Stephen B Soumerai. Systematic review: the relationship between clinical experience and quality of health care. *Annals of Internal medicine*, 142(4):260–273, 2005.

[121] Fredrik Carlsen and Jostein Grytten. More physicians: improved availability or induced demand? *Health Economics*, 7(6):495–508, 1998.

[122] Joseph J Doyle Jr, Steven M Ewer, and Todd H Wagner. Returns to physician human capital: Evidence from patients randomized to physician teams. *Journal of health economics*, 29(6):866–882, 2010.

[123] Jishnu Das and Jeffrey Hammer. Quality of primary care in low-income countries: facts and economics. *Annu. Rev. Econ.*, 6(1):525–553, 2014.

[124] Edward Okeke. Working hard or hardly working: Health worker effort and health outcomes. 2019.

[125] Andrew J Epstein, Sindhu K Srinivas, Sean Nicholson, Jeph Herrin, and David A Asch. Association between physicians' experience after training and maternal obstetrical outcomes: cohort study. *BMJ*, 346, 2013.

[126] William N Evans and Diana S Lien. The benefits of prenatal care: evidence from the pat bus strike. *Journal of Econometrics*, 125(1-2):207–239, 2005.

[127] Gissele Gajate-Garrido. The impact of adequate prenatal care on urban birth outcomes: an analysis in a developing country context. *Economic Development and Cultural Change*, 62(1):95–130, 2013.

[128] Fidel Gonzalez and Santosh Kumar. Prenatal care and birthweight in mexico. *Applied Economics*, 50(10):1156–1170, 2018.

[129] Susan Godlonton and Edward N Okeke. Does a ban on informal health providers save lives? evidence from malawi. *Journal of development economics*, 118:112–132, 2016.

[130] Terhi Johanna Lohela, Robin Clark Nesbitt, Alexander Manu, Linda Vesel, Eunice Okyere, Betty Kirkwood, and Sabine Gabrysch. Competence of health workers in emergency obstetric care: an assessment using clinical vignettes in brong ahafo region, ghana. *BMJ open*, 6(6), 2016.

[131] Edward Okeke, Peter Glick, Amalavoyal Chari, Isa Sadeeq Abubakar, Emma Pitchforth, Josephine Exley, Usman Bashir, Kun Gu, and Obinna Onwujekwe. The effect of increasing the supply of skilled health providers on pregnancy and birth outcomes: evidence from the midwives service scheme in nigeria. *BMC Health Services Research*, 16(1):1–9, 2016.

[132] Leiyu Shi, James Macinko, Barbara Starfield, Jiahong Xu, Jerrilynn Regan, Robert Politzer, and John Wulu. Primary care, infant mortality, and low birth weight in the states of the usa. *Journal of Epidemiology & Community Health*, 58(5):374–380, 2004.

[133] Mengping Zhou, Luwen Zhang, Nan Hu, and Li Kuang. Association of primary care physician supply with maternal and child health in china: a national panel dataset, 2012–2017. *BMC Public Health*, 20(1):1–10, 2020.

[134] Luiz Felipe Campos Fontes, Otavio Canozzi Conceição, and Paulo de Andrade Jacinto. Evaluating the impact of physicians' provision on primary healthcare: Evidence from brazil's more doctors program. *Health economics*, 27(8):1284–1299, 2018.

[135] Thomas Hone, Timothy Powell-Jackson, Leonor Maria Pacheco Santos, Ricardo de Sousa Soares, Felipe Proenço de Oliveira, Mauro Niskier Sanchez, Matthew Harris, Felipe de Oliveira de Souza Santos, and Christopher Millett. Impact of the programa mais médicos (more doctors programme) on primary care doctor supply and amenable mortality: quasi-experimental study of 5565 brazilian municipalities. *BMC health services research*, 20(1):1–11, 2020.

[136] Enlinson Mattos and Debora Mazetto. Assessing the impact of more doctors' program on healthcare indicators in brazil. *World Development*, 123:104617, 2019.

[137] Ece A Özçelik, Adriano Massuda, Margaret McConnell, and Marcia C Castro. Impact of brazil's more doctors program on hospitalizations for primary care sensitive cardiovascular conditions. *SSM-population health*, 12:100695, 2020.

[138] Bladimir Carrillo and Jose Feres. Provider supply, utilization, and infant health: evidence from a physician distribution policy. *American Economic Journal: Economic Policy*, 11(3):156–96, 2019.

[139] Martin Foureaux Koppensteiner and Marco Manacorda. Violence and birth outcomes: Evidence from homicides in brazil. *Journal of Development Economics*, 119:16–33, 2016.

[140] Ministry of Health. *Consolidação do Sistema de Informações sobre Nascidos Vivos - 2011*. Coordenação Geral de Informações e Análise Epidemiológica, 2011.

[141] Sara MA Donahue, Ken P Kleinman, Matthew W Gillman, and Emily Oken. Trends in birth weight and gestational length among singleton term births in the united states: 1990–2005. *Obstetrics and gynecology*, 115(2 Pt 1):357, 2010.

[142] WHO. Low birthweight: country, regional and global estimates. 2004.

[143] Kamila Mistry, Kathryn R Fingar, and Anne Elixhauser. Statistical brief# 211. *Obstetrics & Gynecology*, 121(4):904–7, 2013.

[144] OECD. Caesarean sections (indicator). 2020.

[145] Ann K Blanc and Tessa Wardlaw. Monitoring low birth weight: an evaluation of international estimates and an updated estimation procedure. *Bulletin of the World Health Organization*, 83:178–185d, 2005.

[146] Jan Ties Boerma, KI Weinstein, Shea Oscar Rutstein, and A Elisabeth Sommerfelt. Data on birth weight in developing countries: can surveys help? *Bulletin of the World Health Organization*, 74(2):209, 1996.

[147] Andrew AR Channon, Sabu S Padmadas, and John W McDonald. Measuring birth weight in developing countries: does the method of reporting in retrospective surveys matter? *Maternal and child health journal*, 15(1):12–18, 2011.

[148] Hilary W Hoynes and Diane Whitmore Schanzenbach. Consumption responses to in-kind transfers: Evidence from the introduction of the food stamp program. *American Economic Journal: Applied Economics*, 1(4):109–39, 2009.

[149] Antônio Augusto Moura da Silva, Leopoldo Muniz da Silva, Marco Antonio Barbieri, Heloísa Bettiol, Luciana Mendes de Carvalho, Valdinar Sousa Ribeiro, and Marcelo Zubaran Goldani. The epidemiologic paradox of low birth weight in brazil. *Revista de saude publica*, 44:767–775, 2010.

[150] Marina Clarissa Barros de Melo Lima, Genyklea Silva de Oliveira, Clélia de Oliveira Lyra, Angelo Giuseppe Roncalli, and Maria Angela Fernandes Ferreira. The spatial inequality of low birth weight in brazil. *Ciencia & saude coletiva*, 18(8):2443–2452, 2013.

[151] David L Lee, Justin McCrary, Marcelo J Moreira, and Jack Porter. Valid t-ratio inference for iv. *arXiv preprint arXiv:2010.05058*, 2020.