The influence of reproductive information quality on the probability of unplanned and unwanted pregnancies in Brazil

ABSTRACT

Background: Unplanned pregnancies are a significant risk factor for inadequate use of prenatal care, and unplanned newborns are prone to having low birth weight. Women with unplanned pregnancies have a higher probability of reporting medical problems before and during pregnancy. In fact, the wellbeing of the entire household may be affected. Moreover, unplanned pregnancies have been associated with a higher social burden on taxpayers. Methods: The paper uses propensity score matching approaches to estimate the effect of having correct fertility information on the probability of having unplanned pregnancies. The data was collected from a nationally representative sample of Brazilian women between the ages of 15 and 49 years. Results: Only 26% of pregnant women have the correct information about fertility levels over the menstrual cycle. Women endowed with correct information are 12% less likely to have unwanted pregnancies and 24% less likely to have unplanned pregnancies. Conclusions: Basic fertility knowledge is an important predictor of unplanned pregnancies in Brazil, but only a small share of Brazilian women have this knowledge. More optimistically, offering access to basic fertility information to women of childbearing age can significantly decrease the instances of unplanned pregnancies, thus generating significant benefits to public health and social security systems.

RESUMO

Introdução: A gravidez não planejada é um fator de risco significativo para uso inadequado do cuidado pré-natal, e os recém-nascidos não planejados são propensos a ter baixo peso ao nascer. Mulheres com gravidez não planejada têm maior probabilidade de relatar problemas médicos antes e durante a gravidez. De fato, o bem-estar de toda a família pode ser afetado. Além disso, gravidezes não planejadas têm sido associadas a maior ônus social para os contribuintes. Métodos: O artigo usa abordagens de Propensity Score Matching para estimar o efeito de ter informações corretas de fertilidade sobre a probabilidade de ter gravidezes não planejadas. Os dados foram coletados de uma amostra nacionalmente representativa de mulheres brasileiras com idades entre 15 e 49 anos. Resultados: Apenas 26% das mulheres grávidas têm informações corretas sobre os níveis de fertilidade ao longo do ciclo menstrual. Mulheres com informações corretas têm 12% menos chances de ter uma gravidez indesejada e 24% menos probabilidade de ter uma gravidez não planejada. Conclusões: O conhecimento básico sobre fertilidade é um importante preditor de gravidez não planejada no Brasil, no entanto apenas uma pequena parcela das mulheres brasileiras tem esse conhecimento. De forma mais otimista, oferecer acesso a informações básicas sobre fertilidade para mulheres em idade fértil pode diminuir significativamente os casos de gravidez não planejada, gerando benefícios significativos para os sistemas de saúde pública e de seguridade social.
Introduction

Public health scholars, policymakers, and social scientists have long been interested in the effects of family structure on wellbeing. This paper focuses on unplanned and unwanted pregnancies. There is overwhelming empirical evidence of the unfavorable social and health outcomes of unplanned and unwanted pregnancies. These problems may impact the wellbeing of unplanned/unwanted children, their mothers, their families, and society as a whole.

While a broad literature review is beyond the scope of this paper, it is noteworthy to discuss a few findings that are important to debates around public health policy. For instance, unwanted children are more likely to live in a single-parent family as well as to live in poverty, receive welfare, and die as an infant (Gruber et al., 1999). Women having unwanted children receive many thousands of dollars in cash, food, housing, and medical assistance (Cook et al., 1999). Unplanned pregnancies are also associated with inadequate prenatal care (Delgado-Rodríguez et al., 1997; Eggleston, 2000) and low birth weight (Eggleston et al., 2001). Unplanned pregnancies can also negatively affect the health outcomes of siblings of unplanned children (Lordan & Frijters, 2013). Moreover, if unplanned pregnancies are positively correlated with unmarried pregnancies and abortions, then unplanned pregnancies may lead to higher crime rates. For instance, research has found that unmarried pregnancies are positively associated with the rates of murder and property crime (Kendall & Tamura, 2010) and that the legalization of abortion helped to reduce crime rates in the U.S. (Levitt, 2004).

The above literature reveals a demand for studies about the drivers of unplanned and unwanted pregnancies to inform public health policy. In response, research has found a number of socioeconomic characteristics that are correlated with unplanned/unwanted pregnancies (we review this literature below). We add to this discussion by examining the effects of information quality on the probability of unplanned/unwanted pregnancies. In particular, we investigate whether the probability of having unplanned/unwanted pregnancies is associated with the quality of information women have about their reproductive system, i.e. basic knowledge about the fertile period of the menstrual cycle.

Methods

Population, sampling, and definition of variables

This work uses data from a nationally representative survey implemented by the Brazilian Ministry of Health. The National Survey of Demographics and Health of Children and Women (translated by the author from the original title in Portuguese: “Pesquisa Nacional de Demografia e Saúde da Criança e da Mulher”) randomly sampled 14,617 households across the country. In these households, 15,575 women between the ages of 15 and 49 were interviewed. The stratified sampling design consisted of two steps. First, random census blocks were selected. Second, from each block, twelve households were randomly chosen for an interview. The data were collected between March 11, 2006, and March 5, 2007, and is available to the public at the Brazilian Ministry of Health’s website. The dataset and its documentation can be downloaded from http://bvsms.saude.gov.br/bvs/pnds/banco_dados.php. Figure 1 shows a map of Brazil with its five census regions and the number of sampled households in each region (in parenthesis).

The survey collected demographic variables and various questions regarding the availability of information and pregnancy status. Our focus is on unplanned/unwanted pregnancies. The survey asked women if they were pregnant. If yes, a follow-up question asked her to choose one out of three options: i) wanted to be pregnant at that time; ii) wanted to wait more; iii) did not want to have a child. Responses allowed for the identification of desired pregnancies (option i), miss-timed pregnancies (option ii), and unwanted pregnancies (option iii).

The sample contains 9,280 women who were not pregnant, and 412 who were pregnant. Amongst pregnant women, 156 pregnancies were desired (option i), 166 were miss-timed (option ii), and 90 were unwanted (option iii). Using this information, we construct two binary indicators: Unplanned Pregnancy equals 1 for miss-timed or unwanted pregnancies, 0 otherwise; Unwanted Pregnancies equals 1 for women that did not want to have a child, 0 otherwise.

The paper aims to establish a link between information quality and family planning. Specifically, we are interested in the effect of information quality on the probability of having unplanned/unwanted pregnancies. Obtaining objective measures of information quality is a difficult task. There are not...
only challenges related to quantifying information quality but also challenges related to the multidimensional nature of information sets (i.e., what type of information is relevant).

Our data allow us to construct a variable that can qualify information about a biological characteristic of the reproductive system: the time a woman is fertile. Women were asked the following question: “What is the time of the menstrual cycle, that is, between the beginning of a period and the beginning of another, when a woman has a greater chance of getting pregnant?”. Respondents were instructed to choose one of the following options: a) during the period; b) just after the period is over; c) in the middle of the menstrual cycle; d) just before the period begins; e) at any time during the menstrual cycle; f) other; g) I do not know.

In general, a woman is most fertile at the time of ovulation, which usually occurs 12-14 days before the next period starts (Fehring et al., 2006). A woman’s fertile window normally consists of the 5 days before ovulation and the day of ovulation itself. This happens because spermatozoa can survive for several days (Lynch et al., 2006). Although cycle lengths vary, we consider answer c) to be the correct answer as it applies to the average woman. We construct a binary indicator for the quality of a woman’s reproductive information set, namely Correct Information, that equals 1 if the respondent answered c), 0 otherwise.

The survey also allows us to develop proxy variables for the size of a woman's information set, in other words, general information availability. Three variables capture the amount of contact with communication media: an indicator for reading newspapers (or magazines) almost every day; an indicator for listening to the radio almost every day; and an indicator for watching television almost every day. We also capture knowledge about contraceptive methods with binary variables that indicate whether the woman knows about contraceptive pills, condoms, and day-after pills (also known as next-day pills or plan-B pills). In addition, the survey collected demographic information, namely: age, marital status, and race, and whether or not the respondent’s residence is located in a rural area.

**Statistical analysis**

The paper’s goal is to estimate the causal effect of information quality on the probability of unplanned/unwanted pregnancies. Our statistical analysis is based on the literature of treatment effect estimation, specifically propensity scores methods (Heckman et al., 1997, 1998). We rely on the perspective that women who chose c) to answer the fertility question were “treated” by society (or the health system) and received correct information about their reproductive system. A challenge for the identification of the effect of information quality on the probability of unplanned/unwanted pregnancy is that women that have the correct information about fertility might be quite different from women that have incorrect information. Therefore, simple comparisons of rates of unplanned/unwanted pregnancies between the two groups of women may lead to biased estimates of the effect of information quality on pregnancy because women in the control group (incorrect information) are not necessarily a good benchmark for those in the treatment group (correct information). This happens when women with incorrect information are systematically different from women with the correct information and, as a result, other characteristics of the control group may be driving pregnancy outcomes.

To overcome this challenge, our statistical analysis relies on propensity score matching approaches to estimate the average treatment effect on the treated (ATET), i.e. the average difference between the unplanned/unwanted pregnancy probability of women with the correct information and the unplanned/unwanted pregnancy probability of the same women if they had not been correctly informed. As opposed to using all available data about women that do not have the correct information, we match women in the treatment group with a control group composed of women that have incorrect information and are likely similar to those women with the correct information.

The paper uses propensity scores to find good matches for women in the “treatment” group and benchmark them against women in the “control” group. The propensity score matching method is implemented in two steps. First, we estimate a Logit regression model for the propensity of pregnant women to receive the correct information. We use our set of controls as factors that affect the likelihood of receiving the correct information to calculate propensity scores. These propensity scores represent an estimate of the probability of receiving treatment. Second, these scores are used to find matches for treated women (i.e. to construct control groups). The idea is that the propensity scores can be used to construct better control groups. In other words, we rely on control groups composed of women that, except for the fact that they have incorrect fertility information, are otherwise similar to those in the treatment group. This allows us to “isolate” the effect of information quality. We use three criteria to construct control groups: nearest neighbor matching, radius matching, and kernel matching. These criteria are summarized below.

**Nearest Neighbor**: In the nearest neighbor method, the match for a treated woman is a control woman that minimizes the difference between propensity $i$ and propensity $j$.

**Radius**: In the radius method, woman $i$’s match are all women $j$ such that the difference between propensity $i$ and $j$ is less than or equal to a constant $r$ (we use $r = 0.01$). Note that a lower $r$ increases the quality of the treatment-control match, but it may decrease the number of women in the control group. The research may also lose observations from
the treatment group if a match with propensity difference less than \( r \) is not found.

**Kernel** In the kernel method, each treated woman \( i \) is matched with all control women \( j \), with each control woman being assigned a weight that is inversely proportional to the difference between \( i \)'s propensity and \( j \)'s propensity.

Once the control group is defined, one can use the following formula to calculate the ATET:

\[
\text{ATET} = \frac{1}{n_i} \sum_i Y_i - \sum_j w(i,j)Y_j
\]

where \( n_i \) is the number of women that received correct information being considered in the matching approach, \( Y_i \) is the unplanned pregnancy outcome of woman \( i \) (treated), \( Y_j \) is the unplanned pregnancy outcome of \( j \)'s control group, and \( w(i,j) \) are the weights used in the kernel matching (all control women receive the same weight in the nearest neighbor matching and radius matching).

**Results**

**Socio-demographic variables, information, and pregnancy**

Table 1 shows the summary statistics of our data. The table shows that 62% of Brazilian pregnancies are unplanned and 22% are unwanted. To put this number in perspective, half of pregnancies in America are unplanned (Mulligan, 2015), while in Ghana 70% are unplanned and 31% are unwanted (Eliason et al., 2014).

Table 1. Description of the sample of pregnant women (\( n = 412 \))

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pregnancy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unwanted</td>
<td>0.218</td>
<td>0.414</td>
</tr>
<tr>
<td>Unplanned</td>
<td>0.621</td>
<td>0.486</td>
</tr>
<tr>
<td>Quality of Information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct Information</td>
<td>0.257</td>
<td>0.438</td>
</tr>
<tr>
<td>Availability of Information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reads Newspaper</td>
<td>0.192</td>
<td>0.394</td>
</tr>
<tr>
<td>Listens to the Radio</td>
<td>0.604</td>
<td>0.490</td>
</tr>
<tr>
<td>Watches TV</td>
<td>0.898</td>
<td>0.303</td>
</tr>
<tr>
<td>Knows Contraceptive Pills</td>
<td>0.944</td>
<td>0.23</td>
</tr>
<tr>
<td>Knows Condoms</td>
<td>0.883</td>
<td>0.321</td>
</tr>
<tr>
<td>Knows Day-after Pills</td>
<td>0.039</td>
<td>0.193</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>25.43</td>
<td>6.182</td>
</tr>
<tr>
<td>Married</td>
<td>0.830</td>
<td>0.376</td>
</tr>
<tr>
<td>White</td>
<td>0.333</td>
<td>0.472</td>
</tr>
<tr>
<td>Rural</td>
<td>0.330</td>
<td>0.471</td>
</tr>
</tbody>
</table>

We identify that approximately 26% of pregnant women had the correct information about fertility and reproduction. With respect to sources of information, we find that only 19% of pregnant women read the newspapers, 60% listen to the radio, while the 90% watch television. Contraceptive pills and condom are known to most women (94% and 88%, respectively), while day-after pills are only known to 4% of the interviewed women. Finally, with respect to socio-demographic characteristics, the average age is 25.4 (years), 83% are married, 33% are white, and 33% live in the rural regions.

**Matching analysis**

Results of the propensity score first stage regression are presented in Table 2. The right-hand side variables of the Logit regression capture the respondent’s availability of information and demographic characteristics. The table reports the marginal effects of these variables on the probability of having the correct information. Women that read newspapers are 13% more likely to have the correct information about fertility and reproduction (p-value < 0.05). Watching TV increases the likelihood of having the correct information by 14% (p-value < 0.05). Women that know about condoms are 17% more likely to have the correct information (p-value < 0.01). In fact, knowledge about condoms is the strongest predictor of the likelihood of having the correct information, both in the magnitude of the effect and statistical significance. The probability of having the correct information increases with experience. Results show that for one additional year of age, the probability of having the correct information increases by 0.7%. White women are 8% more likely to have the correct knowledge about reproduction than women from other ethnic backgrounds. The remaining variables were not statistically significant predictors of the probability of having the correct information.

Next, we use the three methods discussed above to select matches (or control groups) for treated women and estimate ATETs. Results are presented in Table 3. The second column shows that, out of the 412 pregnant women in the sample, 106 chose c) to answer the fertility question and are, therefore, considered to have been “treated”, i.e. they received the correct information about fertility. Note that the radius approach only considers 105 women in the treatment group. This happens because for one woman it was not possible to find a match with a propensity score difference less than \( r \). The number of women selected for the control group is displayed in the third column and varies as a function of the matching method. Note that the nearest neighbor method selects only 101 women for the control group. This happens because a woman can be a control observation (i.e. the nearest neighbor) to more than one woman in the treatment group. Also, note that the kernel approach uses all 306 women that do not have the correct information as control; however, it assigns low weights for poor matches.
The influence of reproductive information quality on the probability of unplanned and unwanted pregnancies in Brazil

A influência da qualidade da informação reprodutiva na probabilidade de gravidez não planejada e indesejada no Brasil

J Bras Econ Saúde 2019;11(1):3-9

Table 2. Logit estimation of the propensity of having correct information

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability of Information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reads Newspaper</td>
<td>0.131**</td>
<td>(0.060)</td>
</tr>
<tr>
<td>Listens to the Radio</td>
<td>-0.014</td>
<td>(0.045)</td>
</tr>
<tr>
<td>Watches TV</td>
<td>0.136**</td>
<td>(0.057)</td>
</tr>
<tr>
<td>Knows Contraceptive Pills</td>
<td>-0.044</td>
<td>(0.111)</td>
</tr>
<tr>
<td>Knows Condoms</td>
<td>0.165***</td>
<td>(0.050)</td>
</tr>
<tr>
<td>Knows Day-after Pills</td>
<td>-0.022</td>
<td>(0.103)</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.007**</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Married</td>
<td>0.045</td>
<td>(0.057)</td>
</tr>
<tr>
<td>White</td>
<td>0.079*</td>
<td>(0.047)</td>
</tr>
<tr>
<td>Rural</td>
<td>-0.029</td>
<td>(0.047)</td>
</tr>
</tbody>
</table>

Marginal effects of a logistic regression. Standard errors are in parenthesis. * P-value < 0.10, ** P-value < 0.05, *** P-value < 0.01.

The fourth column shows the estimates of ATET. The upper panel reports the effect of information quality on the probability of unwanted pregnancies, while the lower panel reports the effect of the probability of unplanned pregnancies. The ATET estimates for unwanted pregnancies vary from -0.116 (kernel) to -0.159 (nearest neighbor), and the t-statistics show that all estimates are statistically significant (p-value < 0.05). Using the more conservative estimates from the radius and kernel matching (that also include larger control groups), we find ATET estimates to be around -0.12. The lower panel of the table examines the relationship between information quality and unplanned pregnancies. This panel shows that all ATET estimates vary from -0.236 (nearest neighbor) to -0.252 (kernel), and t-statistics show strong statistical support for all three estimates (p-value < 0.01).

Discussion

Main findings of this study

This study finds that Brazilian women with the correct information about fertility levels over the menstrual cycle have a lower probability of unwanted and unplanned pregnancies. We estimate that basic fertility knowledge (i.e. that the middle of the menstrual cycle is the most fertile period) decreases the probability of having an unwanted pregnancy by approximately 12%, and decreases the probability of having an unplanned pregnancy by approximately 24%. This finding stresses the vital role that basic health education has as a public health policy instrument in developing countries.

Our study reveals that only 26% of pregnant women in Brazil have correct knowledge about fertility. Our Logit model estimates indicate that women that read the newspapers and watch TV regularly are 13% more likely to have the correct fertility information. Knowledge of condoms is the only contraceptive method knowledge that correlated with having the correct fertility information. Our results also suggest that the correct fertility knowledge increases with experience and that Brazil may suffer from an ethnic bias regarding fertility knowledge as white women are 8% more likely to have the correct fertility information.

What is already known on this topic

Unplanned pregnancies are a significant risk factor for inadequate use of prenatal care (Delgado-Rodríguez et al., 1997; Eggleston, 2000), and unplanned newborns are prone to have low birth weight (Eggleston et al., 2001). When
comparing to women with planned pregnancies, women with unplanned pregnancies have a higher probability of being either too young or too old and of reporting more medical problems before and during pregnancy (Bitto et al., 1997).

Unplanned pregnancies may impose a number of negative externalities. For instance, they may decrease the wellbeing of families (Lordan & Frijters, 2013), raise the social burden on taxpayers (Gruber et al., 1999; Cook et al., 1999), and increase crime rates (Levitt, 2004; Kendall & Tamura, 2010). These effects may be especially acute in developing countries like Brazil where abortion is illegal. In fact, in most countries of South America and Africa, abortion is restricted to cases of maternal risk, mental health, or other health complications (United Nations, 2014).

Several papers in the medical, family planning, and public health literature have examined the determinants of unplanned pregnancies. These studies usually associate unplanned pregnancies with socio-economic characteristics like age, marital status, race, education, wealth, knowledge about contraceptive methods (Williams, 1991; Mbizvo et al., 1997; Eggleston, 1999; Font-Ribera et al., 2008; Ikamari et al., 2013) or high-risk sexual behavior (Song & Ji, 2010). Results are mixed but, in general, young, single, black, poor, and uneducated women are more likely to have unplanned pregnancies.

This study is also related to a large body of literature that examines the role that information plays in different health-care issues. Works in this area include investigations about what information patients use to determine their demand for health services (Witt, 2008), the effects of patients’ information on physicians’ treatment recommendations (Xie et al., 2006), the effectiveness of information policies on teenagers’ cannabis use and heavy drinking (Etilé, 2006), and the relationship between information asymmetry and the efficiency of health-care markets (Blomqvist, 1991; Hirth, 1999), to name a few.

**What this study adds**
Education for girls, economic opportunities for women, and family planning services are successful instruments to improve health and wellbeing (Rao & Samarth, 2010). While various socioeconomic factors can affect unplanned pregnancies, this study finds that basic fertility knowledge is an important predictor of unplanned/unwanted pregnancies in Brazil. The study reveals that Brazilian women lack basic reproductive knowledge and this can have negative health consequences for Brazilian families. More optimistically, the study also reveals that simply having access to basic fertility and reproductive information can significantly decrease the probability of unwanted and unplanned pregnancies.

The conclusions of the study are obtained from a propensity score matching analysis. This statistical method offers an approach to control for systematic differences between women with and without the correct information, and serve as an alternative to a more conventional multivariate Logit regression of unwanted pregnancy on correct information and control variables. In fact, the Logit estimate of the effect of correct information is -0.096, or 9% decrease (p-value < 0.05), underestimating the impacts of instruments that offer basic fertility knowledge to women aged 15-49 as a public health policy. Our study suggests that heuristic learning processes in developing countries may lead to poor health outcomes, highlighting the importance of having mechanisms to deliver accurate health-related information to women of fertile age.

**Limitations of this study**
While the study reports estimates that can inform public health policy, we acknowledge the general challenges related to the conceptualization of unplanned and unwanted pregnancies. The definitions used in this study (and in most of the literature) rely on survey data about women’s pregnancy intentions at the time of their response and do not accommodate the dynamic nature of childbearing decisions (Macleod, 2015). The benefits from health policy informed by the results of this study depends on the rationalization of women’s choices such that, when empowered with correct fertility information, women can make fertility choices that have positive health effects.

**Disclaimer statement**
The findings, interpretations, and conclusions expressed in this paper are entirely those of the author. They do not necessarily reflect the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work.

**Acknowledgments**
This study benefited from insights of participants of the seminar series of the Department of Resource Economics and Environmental Sociology, University of Alberta. The author is solely responsible for any omissions or deficiencies.

**References**
The influence of reproductive information quality on the probability of unplanned and unwanted pregnancies in Brazil

A influência da qualidade da informação reprodutiva na probabilidade de gravidez não planejada e indesejada no Brasil

J Bras Econ Saúde 2019;11(1):3-9


