ISSN 1755-5361

University of Essex

Department of Economics

Discussion Paper Series

No. 696 June 2011

Health Information and Health Outcomes: An Application of the Regression Discontinuity Design to the 1995 UK Contraceptive Pill Scare Case

Emilia Del Bono, Marco Francesconi and Nicky Best

Health Information and Health Outcomes: An Application of the Regression Discontinuity Design to the 1995 UK Contraceptive Pill Scare Case*

Emilia Del Bono[1], **Marco Francesconi**[2], **Nicky Best**[3]

June 2011

Abstract

This paper provides a general formulation of the regression discontinuity (RD) design and shows its general applicability to many epidemiological problems. It then applies the RD method to estimate the effects of the 1995 pill scare in the UK, using individual birth records and aggregate monthly statistics. The results show that, following the announcement of the health warning on the "third generation" pill, conception rates increased by about 7%, with a 9% increase in abortion rates and a 6-7% rise in birth rates. No effect was found on still births, very low birth weight, sex ratios, or average birth weight. There is evidence of a slight increase in the rates of low birth weight births and multiple births and of a considerable reduction in the rate of births with congenital anomalies. Heterogeneity by mother's age and social class is very pronounced, with most of the effects being experienced by women aged less than 25 and of lower socioeconomic status.

Keywords: Regression discontinuity, contraception, fertility, birth weight, health warnings, effect of the media

JEL Classification: C21, D8, I12, J13

* We are grateful to John Ermisch, Stephen Jenkins, Cheti Nicoletti, Steve Pudney, Jewel Thomas, and seminar participants at the BIAS Workshop at Imperial College London and at the MISOC workshop at University of Essex for helpful comments and suggestions. Emilia Del Bono acknowledges the support from the Economic and Social Research Council (ESRC) through the Research Centre on Micro-Social Change (MiSoC) (award no. RES-518-28-001). Nicky Best acknowledges the support of the ESRC through the National Centre for Research Methods BIAS node (award no. RES-576-25-0015).

[1] Corresponding author. Institute for Social and Economic Research, University of Essex, Colchester, Essex, CO4 3SQ, UK; edelbono@essex.ac.uk

[2] Department of Economics, University of Essex; mfranc@essex.ac.uk

[3] Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London; n.best@imperial.ac.uk

Introduction

Recent research in epidemiology has emphasized both the importance of, and the need for, causal inference (Susser, 1991; Robins, Hernán, and Brumback, 2000; Maldonado and Greenland, 2002; Rothman and Greenland, 2005). Many of the recent methodological developments have provided support for the focus on isolating factors that can be seen as causes of a specific disease state or health outcome (Galea, Riddle, and Kaplan, 2010). This study proposes an alternative evaluation method, which has not yet been widely applied to epidemiological research, and illustrates it with a new substantive application.

The method is the regression discontinuity (RD) design. Originally introduced half a century ago by psychologists with applications to education (Thislethwaite and Campbell, 1960), the RD design initially received only scant attention by evaluation research methodologists (Cook and Campbell, 1979; Trochim, 1984). Recently, however, it has become widely popular among econometricians and empirical economists (Angrist and Krueger, 1999; Hahn, Todd, and van der Klaauw, 2001; Imbens and Lemieux, 2008), and it has been used in a variety of evaluation exercises, including educational interventions (Angrist and Lavy, 1999; van der Klaauw, 2002; Jacobs and Lefgren, 2004), disability insurance reforms (Chen and van der Klaauw, 2008), legislation aimed at reducing air pollution (Chay and Greenstone, 2005), the impact of unionization on establishment closure (DiNardo and Lee, 2004), and incumbency advantage in elections (Lee, 2008). More generally, there is an increasing understanding that observational studies should be carefully designed to approximate randomized experiments (Rubin, 2006a and 2008). This background provides a strong methodological motivation as to why the RD design is likely to be an appealing tool that should be considered more often in epidemiology.

Our application refers to the evaluation of the impact of medical information dissemination on women's pregnancy decisions and child birth outcomes. The specific example in this study is the contraceptive pill scare that originated from the health warning disseminated by the UK Committee on Safety of Medicines on 19 October 1995. The warning, which received massive media coverage (Thomas, 2010), was based on new scientific results according to which combined oral contraceptives containing either gestodene or desogestrel (the 'third generation' pill) were associated with twice the risk of venous thromboembolism

as compared to that associated with older products (Poulter, Chang, Farley, Meirek, and Marmot, 1995; Jick, Jick, Gurewich, Myers, and Vasilakis, 1995; Spitzer, Lewis, Heinemann, Thorogood, and MacRae, 1995) These results were later disproved by new evidence (Farmer, Williams, Simpson, and Nightingale, 2000). But, by that time, their consequences on pregnancies and abortions had already unfolded (Wood, Botting, and Dunnell, 1997).

This pill scare event is likely to be a salient shock to women's choice environment. Our hypothesis is motivated by the evidence that the pill scare might have changed women's contraceptive behaviour, with consequent impact on birth rates and birth outcomes (Allison and Reizon, 1996; Hope, 1996; Wood, Botting, and Dunnell, 1997; Furedi, 1999). Such effects, which are crucial to public health policy, are suitable for analysis within the RD paradigm. Assignment to treatment in the RD design — like that in all observational data — is not random and individuals who receive treatment may systematically differ from those who do not. But, in contrast to conventional observational studies, within the RD design the analyst has specific knowledge of the assignment rule that influences how individuals are assigned to (or selected into) treatment. More specifically, the design requires that there is a known *cut-off* point in treatment assignment or in the probability of treatment receipt as a function of one or more assignment variables, generating a discontinuity in the treatment recipiency rate at that point.

In the pill scare application, the treatment is the UK Committee on the Safety of Medicines' health warning and the cut-off point is the date on which the health warning was made public (19 October 1995). The insight is that we take advantage of knowing this cut-off date to learn about the impact of medical information on conception and birth outcomes for women near the cut-off. Assuming that women who were at an advanced stage of pregnancy before the cut-off date represent a valid no-treatment comparison group for those who were not pregnant yet, the analyst could evaluate the impact of the health warning by relating average outcomes for treated women just after the cut-off date with those of untreated women just before it. That is, under certain comparability conditions, the assignment near the cut-off can be seen as behaving almost as if it were random.

Methods

Statistical analysis

Our outline of the RD design is brief and is meant to illustrate the pill scare application. More general descriptions of the RD method are available elsewhere (van der Klaauw, 2008; Imbens and Lemieux, 2008; Lee and Lemieux, 2010).

Consider the problem of evaluating the causal effect of a binary treatment on an outcome variable, using a random sample of individuals where for each woman i we observe an outcome measure Y_i (e.g., having a child or the child's birth weight) and a binary treatment indicator T_i, equal to one if treatment was received and zero otherwise (e.g., having received information about the pill scare or not). Let $Y_i(1)$ denote the potential outcome given treatment, and $Y_i(0)$ the potential outcome in the absence of treatment (Rubin, 2006b). The causal effect of treatment on woman i is defined as

$$\beta_i = Y_i(1) - Y_i(0) \tag{1}$$

The evaluation problem that arises in determining the effect of T on Y comes from the fact that each woman is either exposed or not exposed to the treatment (i.e., she is either exposed to the health warning or not, respectively) and is never simultaneously observed in both states. Hence (1) cannot be observed directly, but rather, the actual observed outcome is $Y_i = T_i Y_i(1) + (1 - T_i)Y_i(0)$. A common regression model representation for this observed outcome can then be written as

$$Y_i = \alpha + \beta_i T_i + \varepsilon_i, \tag{2}$$

where $Y_i(0) = \alpha + \varepsilon_i$, β_i is as defined in (1) and ε_i is a random error term. In general, comparing average observed outcomes of treated and untreated women using a model such as (2) does not yield a good estimate of the average treatment effect $E[\beta_i]$, however. With a little bit of algebra, it can be shown that

$$\begin{aligned} E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 0] &= E[\beta_i] + \big(E[\varepsilon_i|T_i = 1] - E[\varepsilon_i|T_i = 0]\big) \\ &+ \Pr(T_i = 0) \times \big(E[\beta_i|T_i = 1] - E[\beta_i|T_i = 0]\big). \end{aligned} \tag{3}$$

This equation reveals two potential sources of bias. The first is due to baseline differences between treated and untreated women (reflected by the second term on the right hand side

of (3)) and the second is due to treatment heterogeneity (reflected by the final term on the right hand side of (3)). Both these sources of bias can be eliminated if the treatment assignment is random.

In a *sharp* RD design, women are assigned to treatment only on the basis of a cut-off value of an observed continuous variable x. In our application this variable (also known as the assignment, selection, running, or ratings variable), is represented by time over which women acquire medical information that may affect their pregnancy decisions and outcomes. Women who are pregnant before a distinct cut-off date $\overline{x}$ are placed in the control group ($T_i = 0$), while those who are pregnant after that date are placed in the treatment group ($T_i = 1$), or vice versa. The health warning date (19 October 1995) is our cut-off date $\overline{x}$. Thus, assignment occurs through a known deterministic decision rule according to which

$$T_i = T(x_i) = \mathrm{I}\{x_i \geq \overline{x}\},$$

where $\mathrm{I}\{\cdot\}$ is the indicator function. If, conditional on other observable factors, women close to the cut-off date are comparable, then we may view the design as almost experimental near $\overline{x}$, suggesting that we could evaluate the causal impact of treatment by comparing the average outcomes for women with ratings just above and below the cut-off. Formally, consider the following *local continuity assumption*:

$$E[\varepsilon_i|x] \text{ and } E[\beta_i|x] \text{ are continuous in } x \text{ at } \overline{x},$$

then, assuming that the density of x is positive in a neighborhood of $\overline{x}$ and using $E[Y_i|x]$ as shorthand notation for $E[Y_i|x_i = x]$ and $E[\beta_i|\overline{x}]$ for $E[\beta_i|x_i = \overline{x}]$, we have

$$\begin{aligned} \lim_{x\downarrow\overline{x}} E[Y_i|x] - \lim_{x\uparrow\overline{x}} E[Y_i|x] &= \left(\lim_{x\downarrow\overline{x}} E[\beta_i T_i|x] - \lim_{x\uparrow\overline{x}} E[\beta_i T_i|x]\right) + \left(\lim_{x\downarrow\overline{x}} E[\varepsilon_i|x] - \lim_{x\uparrow\overline{x}} E[\varepsilon_i|x]\right) \\ &= E[\beta_i|\overline{x}], \end{aligned} \tag{4}$$

where "$\downarrow$" and "$\uparrow$" denote that the limit is taken as x approaches $\overline{x}$ from above (or just after 19 October 1995) and from below (or just before the same date), respectively. The RD approach of comparing average outcomes just right and left of the cut-off, therefore, does identify the average treatment effect for individuals close to the discontinuity point. The local continuity assumption formalizes the condition mentioned earlier that women

just after and before the pill scare warning are 'comparable', requiring them to have similar average potential outcomes when receiving treatment and when not. Another way to look at this is to note that the local continuity assumption ensures that the second and third terms in (3) are zero, hence eliminating the potential sources of bias discussed earlier. Identification of the average causal treatment effect is thus achieved in the RD approach assuming only smoothness in expected potential outcomes at the discontinuity without any parametric functional form restrictions.

Clearly, treatment assignment may depend on x in a stochastic manner, rather than deterministically as assumed in the sharp design. Nonetheless, the propensity score function $\Pr(T_i = 1|x)$ continues to have a discontinuity at $\overline{x}$. Instead of a 0-1 step function, the treatment probability as a function of x contains a jump at $\overline{x}$ that is less than unity. This is the *fuzzy* RD design (Campbell, 1969; Imbens and Wooldridge, 2007; van der Klaauw 2008).

A fuzzy design allows for values of the running variable near the cut-off date $\overline{x}$ that can belong to either the treatment or control group. This is similar to having 'no-shows' (treatment group members who do not receive treatment) and 'cross-overs' (control group members who receive the treatment) in randomized experiments. For instance, it could be that some women whose conception date was after 19 October 1995 did not follow the ensuing media discussions and were not informed of the risks of oral contraceptives by friends and family. By contrast, since the conception date is derived from the birth date using a standard gestation length of 38 weeks, it could be that some of the women whose hypothetical conception date falls before the cut-off were in fact exposed to the treatment (if their true gestation length was less than 38 weeks, for example). All such circumstances make no-shows and cross-overs likely in our application.

To see how the discontinuity in the selection rule can identify an average treatment effect of interest in the fuzzy RD case, it is worth noting that if $\beta_i = \beta$ in a neighborhood of $\overline{x}$ (i.e., if the treatment effect is locally constant), and if the same local continuity assumption as in the sharp design holds, then $\lim_{x\downarrow\overline{x}} E[\beta_i T_i|x] - \lim_{x\uparrow\overline{x}} E[\beta_i T_i|x]$ in (4) will be equal to $\beta(\lim_{x\downarrow\overline{x}} E[T_i|x] - \lim_{x\uparrow\overline{x}} E[T_i|x])$ and β is identified by

$$\frac{\lim_{x\downarrow\overline{x}} E[Y_i|x] - \lim_{x\uparrow\overline{x}} E[Y_i|x]}{\lim_{x\downarrow\overline{x}} E[T_i|x] - \lim_{x\uparrow\overline{x}} E[T_i|x]}. \tag{5}$$

If instead the treatment effect is heterogeneous, then, besides the local continuity assumption, we also need to invoke a *local conditional independence assumption*, according to which T_i must be independent of β_i conditional on x near $\overline{x}$. Under these two assumptions, the first term in parentheses on the right-hand side of (4) equals $\lim_{x\downarrow\overline{x}} E[\beta_i|x] \lim_{x\downarrow\overline{x}} E[T_i|x] - \lim_{x\uparrow\overline{x}} E[\beta_i|x] \lim_{x\uparrow\overline{x}} E[T_i|x]$. This implies that the ratio in (5) identifies $E[\beta_i|x = \overline{x}]$, the average treatment effect (ATE) for women with values of x close to $\overline{x}$.

Notice that if women self-select into treatment, the conditional independence assumption is likely to be violated. Assume that $T_i(x)$, woman i's treatment assignment given any x, is a deterministic function that varies across individuals. In this case, under a *local monotonicity assumption* (Imbens and Angrist, 1994), it can be shown that the ratio in (5) identifies a local average treatment effect (LATE) at the cut-off date, given by (Han, Todd, and van der Klaauw, 2001):

$$\lim_{\zeta\downarrow 0} E[\beta_i|T_i(\overline{x}+\zeta) - T_i(\overline{x}-\zeta) = 1, x = \overline{x}].$$

This causal effect represents the average treatment effect of the 'compliers', that is, the subgroup of women whose treatment status would switch from nonrecipient to recipient if their pregnancy date x crossed the cut-off date.

In our application, we do not know which women received information about the risks of using the pill and which women were still unaware of these risks even after the media campaign. While we can reasonably assume that before the cut-off date no woman had received the treatment, it is likely that not all women were aware of the risks of the pill the day after 19 October 1995. In principle, we are in a situation where a *fuzzy* design is appropriate, but as we do not directly observe treatment status T_i we cannot use (5) to derive β. We therefore make the following assumption:

(A.1) $\Pr(T_i = 1|\overline{x}+\Delta_2) - \Pr(T_i = 1|\overline{x}-\Delta_1) = 1.$

That is, by allowing for a discrete interval of time around the cut-off date — when the medical information was presumably passed on to different groups of women — *no* woman is assumed to be treated before $\overline{x}-\Delta_1$, while *all* women are assumed to be treated after $\overline{x}+\Delta_2$.

A second problem with our application, and one which is potentially relevant for any type of RD design, is that because of delays in the onset of fecundity and unobserved variation in gestation, we are not able to precisely distinguish births to women who were exposed to the health shock from births to those who were not in the vicinity of the cut-off date. Using a discrete time interval around the cut-off date helps us to mitigate this problem, however there might still be some measurement error in the running variable x and following (Battistini, Brugiavini, Rettore and Weber (2009) we assume that:

(A.2) Conditional on the true value of x, x^*, the process generating the measurement error is orthogonal to the process governing exposure to medical information;

Such assumptions lead us to the deterministic treatment assignment that characterizes the *sharp* design, and to the case with no measurement error in x, so that we can use (4) to recover β. This will have an ATE interpretation under the following *piece-wise continuity* restriction:

(A.3) $E[\varepsilon_i|x]$ and $E[\beta_i|x]$ are continuous in x at $\overline{x}-\Delta_1$ and $\overline{x}+\Delta_2$.

Empirically, we perform our main analysis by imposing a two-week window around $\overline{x}$. We then experiment with a number of alternative (symmetric and asymmetric) time intervals, Δ_1 and Δ_2, to check for the sensitivity of our results to different specifications of the data generation process.

Data

Birth, still birth and congenital anomaly records were extracted from the national births and stillbirths registers and the National Congenital Anomalies System (NCAS), respectively, held at the UK Small Area Health Statistics Unit (SAHSU). Registration of all births and stillbirths is a legal requirement in the UK, providing national registers with high levels of ascertainment. Since the end of 1992, stillbirths are legally defined as fetal deaths after 24 completed weeks of gestation. Congenital anomalies records were matched to birth records using sex, date of birth, postcode, birth weight and maternal age. Small area socioeconomic indicators derived from the 1991 and 2001 UK censuses were also linked

to birth records using postcode-to-enumeration district link files. For a 10% random subsample, we also have information on maternal social class. Each birth date was converted into a conception date by assuming a 38 weeks gestation length. Multiple births were considered as a single conception for the purposes of calculating the number of conceptions leading to births. All records were aggregated into monthly observations and stratified by mother's age at birth.

We also use a second data source from the Office for National Statistics (ONS). This has only aggregate data, but, in addition to births, it also contains monthly information on abortions, allowing us to analyse total conceptions and distinguish between conceptions ending in a birth and conceptions ending in abortions. Both data sources have information on mother's age at birth, which is aggregated into five different groups (i.e., <20, 20–25, 26–29, 30–35, 36+).

The ONS aggregate data were obtained for the period between April 1993 and March 1998, while the register data were extracted from the SAHSU database for the months between April 1994 and March 1998. For analytical purposes, we divided the sample period into four "cohorts", i.e., yearly intervals centered around the month of October, since this was the month in which the pill scare occurred, allowing us to consider (evenly or unevenly) spaced windows of data on both sides of the date of the announcement in October 1995, as well as in years other than 1995.

We analyse two sets of outcomes. The first captures a notion of *quantity* and is given by (i) daily average number of conceptions; (ii) daily average number of abortions; (iii) daily average number of live births. The first two outcomes can be examined only with the ONS aggregate data, the third can be derived from the individual register data as well. The second set of outcomes, which is derived only from the register data, approximates a notion of *quality* and consists of (i) number of still-born babies per 1,000 births (the denominator here and in the other measures always includes live and still births); (ii) number of babies with congenital anomalies per 1,000 births; (iii) average birth weight; (iv) number of babies with very low birth weight (<1,500 grams) per 1,000 births; (v) number of babies with low birth weight (<2,500 grams) per 1,000 births; (vi) number of babies with high birth weight (>4,500 grams) per 1,000 births; (vii) number of multiple births per 1,000 births; and (viii) sex ratio (number of male babies per 1,000 births).

Estimation

The identification results presented earlier indicate that estimation of treatment effects in the case of an RD design requires estimating boundary points of conditional expectation functions. With a sufficiently large number of conceptions, one could focus on units within a very small time interval around the cut-off date and compare average outcomes for women just left and right of the discontinuity point. As mentioned, however, there are problems in pursuing this strategy in our application. Dating conception accurately is difficult, ovulation may be delayed, conception waits can be non-negligible after stopping pill usage, and the health warning itself might have been perceived differently by different women or not received at all. Consequently, we have to increase the interval around the cut-off date in order to pin down a meaningful effect of the pill scare, although increasing this interval could complicate the interpretation of the effect estimate, especially if the assignment variable becomes correlated to birth outcomes conditional on treatment status (e.g., if there are strong seasonal conception/birth regularities).

Additional assumptions, therefore, are introduced about the relationship between the cut-off date and birth outcomes, allowing us to use more observations and extrapolate trends from above and below the cut-off point more easily. We adopt a parametric specification for the conditional expectations functions, which account for heterogeneity in mother's age at birth and seasonal effects. Specifically, we estimate:

$$g\left\{E(Y_{acm})\right\} = \beta_0 + \beta_1 T + \beta(T\times C) + \gamma_1 A + \gamma_2 C + \gamma_3(A\times C) + \delta_0 f(m) + \delta_1(A\times f(m)), \quad (6)$$

with $Y \sim F$, where $g(\cdot)$ is a generic link function and F is a distributional family. If $g(\cdot)$ is the natural log function and Y is distributed as Poisson, then (6) is estimated as a Poisson regression; when instead $g(\cdot)$ is the logit function and Y is distributed as a Bernoulli distribution function, we estimate logistic regressions, and if the link function is the identity function with Y being distributed as Gaussian, then (6) is a linear probability model.

As before, Y represents the outcome of interest, which now however is allowed to vary according to the age group of the mother, a, conception cohort, c, and month within cohort, m, and T assumes value 0 for all months between April and September and value 1 for all months between November and March for every year in the sample, thus splitting each

cohort into two roughly equal intervals. The term A denotes a set of dummy variables for mother's age group; C represents a set of indicators for conception cohort; $f(m)$ is a polynomial function of month within cohort. This specification explicitly allows Y to vary across calendar time in a way that differs according to mother's age (through γ_3). It also allows for common as well as age-specific seasonal effects (through δ_0 and δ_1, respectively).

The parameter of interest, which identifies the effect of the pill scare, is in the vector β. Taking the period between April 1994 and March 1995 (cohort 1994/95) as the reference conception cohort, the parameter we are interested in measures the size of the change in Y before and after October 1995 (cohort 1995/96) relative to the what observed in the previous year. Looking at the change before and after October 1996 (cohort 1996/97), equation (6) also allows us to test also for the presence of placebo effects, as in standard randomized experiments.

As mentioned, we vary the time interval around the cut-off date. In particular, we estimate (6) using a two-month window around October, the month of the health warning, i.e., using August-September and November-December only and omitting October from the analysis. The same model is then re-estimated over a three-month window (i.e., July through the following January, again omitting October). This fanning out process continues stepwise up to a five-month window. The wider the window, the more likely we are to capture effects on birth outcomes (because of conception waits, fecundity delays and informational issues), but the more likely we are also to introduce 'spurious' variation, such as seasonal patterns in birth rates. We present results for all such different windows.

Finally, the quantitative outcomes, which are count variables, are estimated using weighted Poisson regressions, where the weights are given by the total number of births observed by mother's age group, a, cohort, c, and month, m. The qualitative outcomes, except for birth weight, are instead proportions that are derived from the aggregation of dichotomous measures. To analyse such outcomes, model (6) is estimated using grouped logistic regressions, where the dependent variable is, say, the number of still births or the number of low birth weight babies, over the number of all births (live and still births) in a given $a \times c \times m$ cell. The effects on birth weight (expressed in grams) are obtained from weighted linear least squares regressions, where the weights are defined as before. Standard errors are clustered according to the values of the running variable, i.e., month

within cohort (Lee and Card, 2008). All estimations were performed using STATA 11.0 SE.

Results

"Quantity" outcomes

Before focusing on the regression estimates, we perform a graphical analysis to check whether there is any visible discontinuity in conceptions and other outcomes at the cut-off date. We compute the residuals from age-specific weighted Poisson regressions on conceptions, births and abortions controlling for seasonal time effects, using monthly dummy variables. Figure 1 plots such residuals in the case of all conceptions, along with a vertical line corresponding to October 1995 and two linear fits, one at the left and the other at the right of the cut-off date. The figure documents the existence of an overall decreasing trend in the number of conceptions over the observed period. Negative age-specific trends are particularly strong for women aged 20-25 and women aged 26-29, while those for women 30-35 and >36 are clearly positive. What is important, however, is that we see a clear increase in the raw number of conceptions among all women below age 35 after October 1995. Results on conceptions leading to births and conceptions leading to abortions are similar and, because of space limits, are not reported.

Figure 2 plots the residuals from age-specific weighted Poisson regressions in which we include also cohort dummies. The aim of this exercise is to compare the period before to that after the cut-off month of October in the treatment cohort (1995/96), as well as in the previous (1994/95) and following (1996/97) cohorts. The figure shows that for all women, except those in the oldest group, the increase in the number of daily conceptions after the month of October occurs only in the treatment cohort. It also shows that the effect was delayed by about 2-3 months, as the figures show an increase in the outcome variable from January 1996 onwards. Again, comparable results on live births and abortions are not displayed.

We now quantify the magnitude of such effects for all outcomes. Table 1 reports the *quantity* estimates on all conceptions, as well as conceptions leading to births and conceptions terminated by abortion. The figures in the table represent the percentage increase in

conceptions, abortions or conceptions leading to a birth which could be attributed to the pill scare.

The pill scare led conceptions to increase significantly between 5.7 and 7.4%, depending on whether we impose a two-month or a five-month window around the cut-off date (panel A, first row). Between 20 and 25% of this increase was due to conceptions terminated by abortion. This implies a substantial rise of between 6.4 to 9.5% in the number of abortions (second row). The remaining increase in conceptions led to an increase in births (third row), which in the five-month window after October rose by 6.9% (95% CI: 3.9%–10.0%). Similar results emerge from the register data (panel B), although in this case the increase in the number of births was slightly smaller and estimated to be 6.0% (95% CI: 4.0%–8.1%) in the five-month window case.

"Quality" outcomes

Table 2 shows the results on birth *quality*, which can be estimated only with the individual register data. All estimates are expressed as odds ratios, except for the case of birth weight where they are in grams.

There is no evidence of a significant effect of the pill scare on the rates of still births, very low birth weight, high birth weight and sex ratio at birth as well as on the average birth weight. The estimates, however, indicate that babies born to women exposed to the pill scare were significantly *less* likely to be affected by congenital anomalies. This effect becomes lower in magnitude as the time window around the cut-off month increases, but the odds are still significantly below one in the ± 5 month case.

The pill scare also led to babies which were *more* likely to be low birth weight and born from multiple pregnancies. It is possible that these two processes are correlated, with the increase in multiple pregnancies resulting in the increase in low birth weight babies. It should be stressed, however, that the effect estimates are relatively small and that they emerge only with short time windows around October. Interestingly, performing the analysis by women's age group, we found evidence of this double effect for women in all age groups except for those aged less than 20. The latter were *more* likely to have a low birth weight child but *less* likely to experience multiple pregnancies.

Placebo test

We checked whether our results are driven by the health warning of October 1995 and not by a discontinuity in the process underlying conceptions at that time of *any* year.

To do so, we looked at whether there is a significant change in outcomes after October in each of the years 1993, 1996 and 1997 compared to after October 1994 (the reference cohort). There were no health warnings or policy changes that might have affected fertility outcomes after such periods. Thus, we expect to observe no change in those years if the 1995 pill scare was responsible for the responses estimated so far. Indeed, the results for 1996 and 1997 (not presented for the sake of brevity) confirm this expectation, showing no significant change in either birth quantity and birth quality outcomes. We did find small increases in conception and birth rates (of up to 2%) after October 1993, although such estimates were highly sensitive to the window around October and often statistically insignificant, and we found no effect on abortion rates.

Response heterogeneity

As revealed by Figures 1 and 2, fertility effects may be heterogenous across women's age groups. There may be also heterogeneity across other important determinants of birth outcomes, such as social class and area of residence.

To assess the extent to which this was indeed the case, we first split women into five age groups and repeated the analysis on each of them separately. Figure 3 displays the results on the quantity outcomes with the May-March ($\pm$5 month) window. Most of the conception effects are concentrated amongst younger women. For example, the pill scare led to an increase in all conceptions ranging between 8.0% (95% CI: 4.6%–11.5%) and 10.6% (95% CI: 5.9%–15.6%) for the groups below 30. For older women, the corresponding increases were smaller: 4.8% (95% CI: 1.7%–8.1%) for those aged 30–35 and 2.8% (95% CI: -2.0%–7.9%) for those aged 36 or above. Abortion rates also increased considerably among younger women, on average between 10.9 and 12.7% for those aged 25 or below and less than 20, respectively. The change in conceptions leading to a birth was thus more evenly spread across all age groups. In terms of the quality outcomes (not shown), we typically found smaller and statistically insignificant differences across age groups. An

interesting exception is on multiple births, which declined among mothers aged less than 20 and increased among older mothers, especially those aged 30–35. This last response was also accompanied by a small increase in the number of low birth weight babies.

ONS releases information on maternal social class for about 10% of birth records each year. We recoded that information into four occupational groups: (i) professional, managerial, and technical (22.4% of the subsample); (ii) skilled non-manual (22.4%); (iii) skilled manual, partly skilled, and unskilled (14.5%); and (iv) out-of-the labour force (40.7%). The pill scare quantity effects emerged primarily amongst women in this last category, who experienced an increase in the birth rate of 10.0% (95% CI: 5.5%–14.8%) in the five months following the warning. For women in skilled non-manual positions, the effect was lower (7.2%, 95% CI: 0.8%–13.9%), while for the other social class groups the effect was not statistically significant. On the other hand, no class gradient was found along birth quality outcomes.

The out-of-the labour force category may identify a highly diverse group of women. To better ascertain the presence of heterogenous responses, we thus stratified our sample on an area-level index of deprivation constructed from the 1991 and 2001 censuses that was linked to birth records using postcode-to-enumeration district link files. The sample was divided into four groups based on quartiles of the deprivation index (Carstairs score), with babies in the first quartile being born to mothers who were living in the most affluent areas and babies in the fourth quartile being born to mothers who were living in the most deprived areas. For conceptions leading to birth, the strongest response was found among women in the latter group, with an increase of 8.3% (95% CI: 6.3%–10.3%). The effects were smaller (between 5.1 and 6.1%), but still statistically significant, among women from the other quartiles. We found no evidence of an impact on birth quality outcomes.

Sensitivity analysis

To validate our results we performed a series of robustness checks. We focussed on quantity outcomes, since most of the treatment effect responses were found for such outcomes. The results are summarised in Table 3.

First, because identification of β rests on piece-wise continuity in the running variable

of both the conditional treatment effect and the unobservable processes underlying our outcomes, we allowed for different time intervals around $\overline{x}$. To check whether the assumption of a delayed response of two weeks (Δ_2 =14 days) was restrictive, we re-estimated (6) after omitting the month of November and not just October, i.e., imposing Δ_2 =44 days, while Δ_1 remained at 17 days. The quantity effects shown in panel A of Table 3 are slightly greater than (but never statistically significantly different from) those reported in Table 1. This suggests that the two-week interval around $\overline{x}$ was somewhat limiting, but the piece-wise continuity assumption does not seem to be violated. Omitting also the month of September from estimation (and, thus, imposing Δ_1 =47 days and Δ_2 =44) did not lead to different results.

So far, we imposed symmetric time windows, from ± 1 to ± 5 months. We therefore checked the sensitivity of our estimates to asymmetric windows, which allow for a longer time to treatment. We thus imposed windows of $-3/+5$ and $-3/+8$ months. The results are in panel B of Table 3. To ease the comparison with our earlier estimates, the table also reports the estimates found with the ± 3 month window reported in Table 1. For all quantity outcomes, the point estimates increase with the 5-month window after the health warning but tend to revert to the symmetric 3-month window estimates when the time window to the right was widens to eight months. For the quality outcomes (not shown), using an asymmetric time window of 5 or 8 months to the right of the cut-off month confirmed virtually all our earlier findings.

Third, several alternative specifications of seasonal effects ($f(m)$ in (6)) were analysed. Specifically, we allowed for a quadratic trend to differ before and after the month of October of each cohort. The estimates in panel C of Table 3 are virtually identical to those we presented in Table 1. We also replaced the quartic function of month within the cohort-year used in the main analysis with different order polynomials. The estimates were invariably very close to those shown in Table 1.

Finally, it is important to quantify the extent to which our RD effect estimates differ from estimates which could be obtained using standard time-series strategies sometimes adopted in the medical literature (Wagner, Soumerai, Zhang, and Ross-Degnan, 2002). To this end, we estimate a before-after model as given in (2) using the following variant of

specification (6):

$$g\{E(Y_{acm})\} = \alpha_0 + \pi T' + \psi_1 A + \psi_2 C + \psi_3 (A \times C) + \lambda_0 f(m) + \lambda_1 (A \times f(m)), \quad (7)$$

where π is now expected to capture the effect of the pill scare, with the new treatment variable, T', being now equal to 1 for all months after October 1995 and zero for all months before. Estimates of π for the quantity outcomes are reported in panel D of Table 3. Two remarks are in order. First, all 'standard' treatment estimates are almost always smaller than their corresponding RD counterparts. Second, the difference is quantitatively important, with 'standard' estimates being 23, 14 and 25% lower than the corresponding RD estimates for conceptions, abortions and births (± 5 window), respectively.

We repeated the same exercise for the birth quality outcomes. The estimates are not reported for the sake of brevity. Two results are worth stressing. First, as in the RD design case, there is evidence of a significant reduction in the rate of births affected by congenital anomalies. But the reduction is estimated to be around 6.7%, approximately 40% less than what we estimated using the RD method. Second, we find no significant effects on low birth weight or on multiple births.

Conclusions

This paper offers two main contributions. The first is methodological. This is one of the first applications of the regression discontinuity evaluation method to an epidemiological issue. It thus offers an overview of the rapidly growing econometric literature in this area and discusses the key assumptions needed to make causal inference in a context where assignment to treatment is a discontinuous function of a known variable.

The second contribution is substantive. For the first time, we apply the RD design to estimate of the effects of the 1995 pill scare in the UK, using individual birth records and aggregate monthly statistics. The results show that, following the announcement of the health warning on the "third generation" pill, conception rates increased by about 7%, with a 9% increase in abortion rates accounting for one quarter of that growth and the remaining three-quarters being accounted for by a 6-7% rise in birth rates.

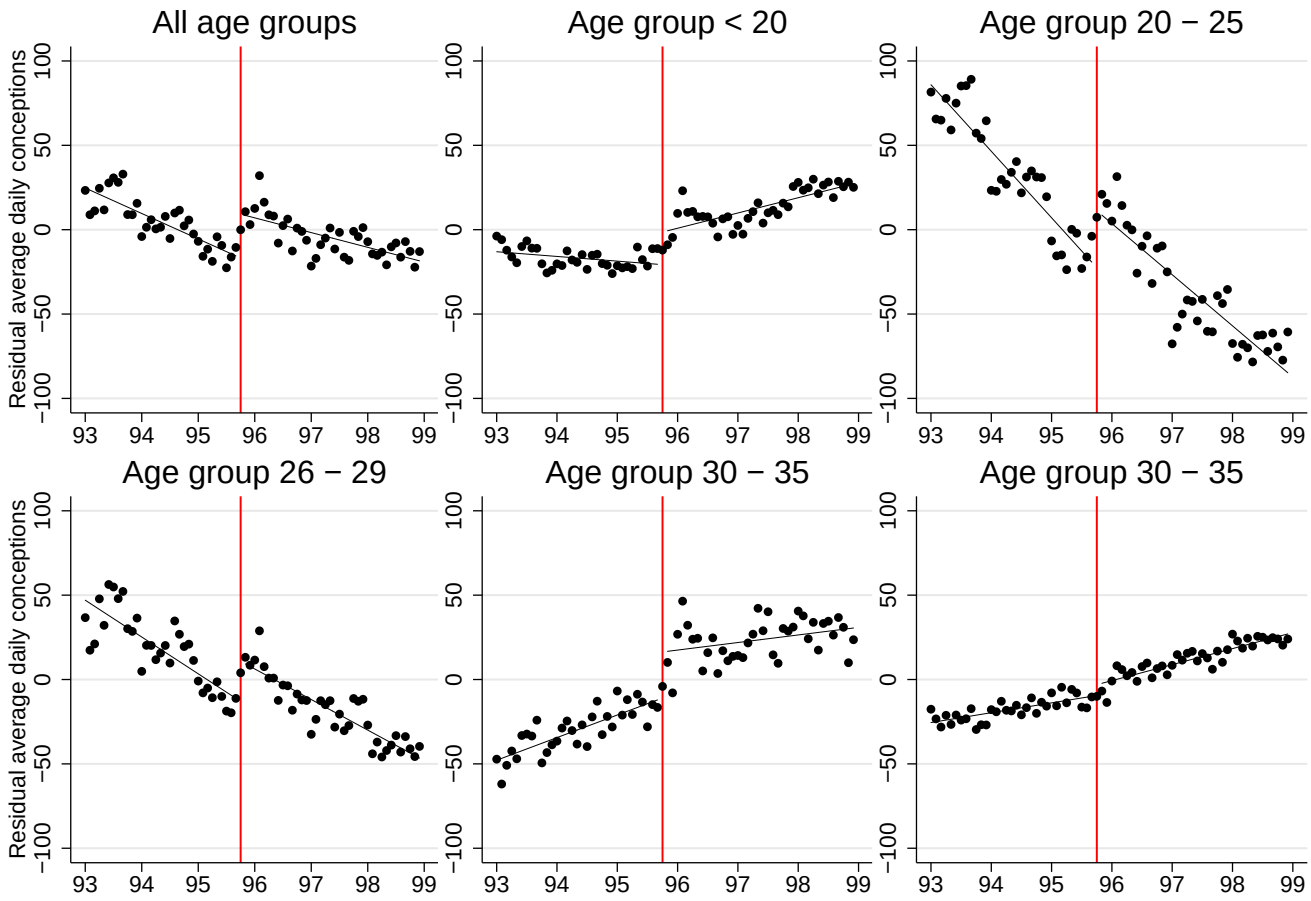
Entirely new is also the analysis of the qualitative outcomes. For these, we found no effect on the rates of still births, very low birth weight births and unequal sex ratios at birth

and on average birth weight births. We found, however, evidence of a slight increase in the rates of low birth weight births and multiple births and a substantial reduction in the rate of births with congenital anomalies. There is also a considerable effect of heterogeneity by mother's age and social class, with most of the quantity effects being experienced by women aged less than 25 and by women with lower socioeconomic status.

The paper's substantive findings stress the relevance of medical information in general and how this is disseminated among (and perceived by) the public at large in particular (Stross and Harlan, 1979; Davey-Smith 2008). These issues are key to public health policy (Pearce, 2004 and 2011; Loewenson, 2008), and therefore measuring their impacts accurately is of paramount importance. The results found using standard analyses of the same problem would lead to up to 30% to 40% downward biased estimates of some of the outcomes. For other outcomes, standard methods would have not been able to detect significant effects at all. The assessment about the effects of the pill scare, therefore, would differ substantially depending on the identification strategy.

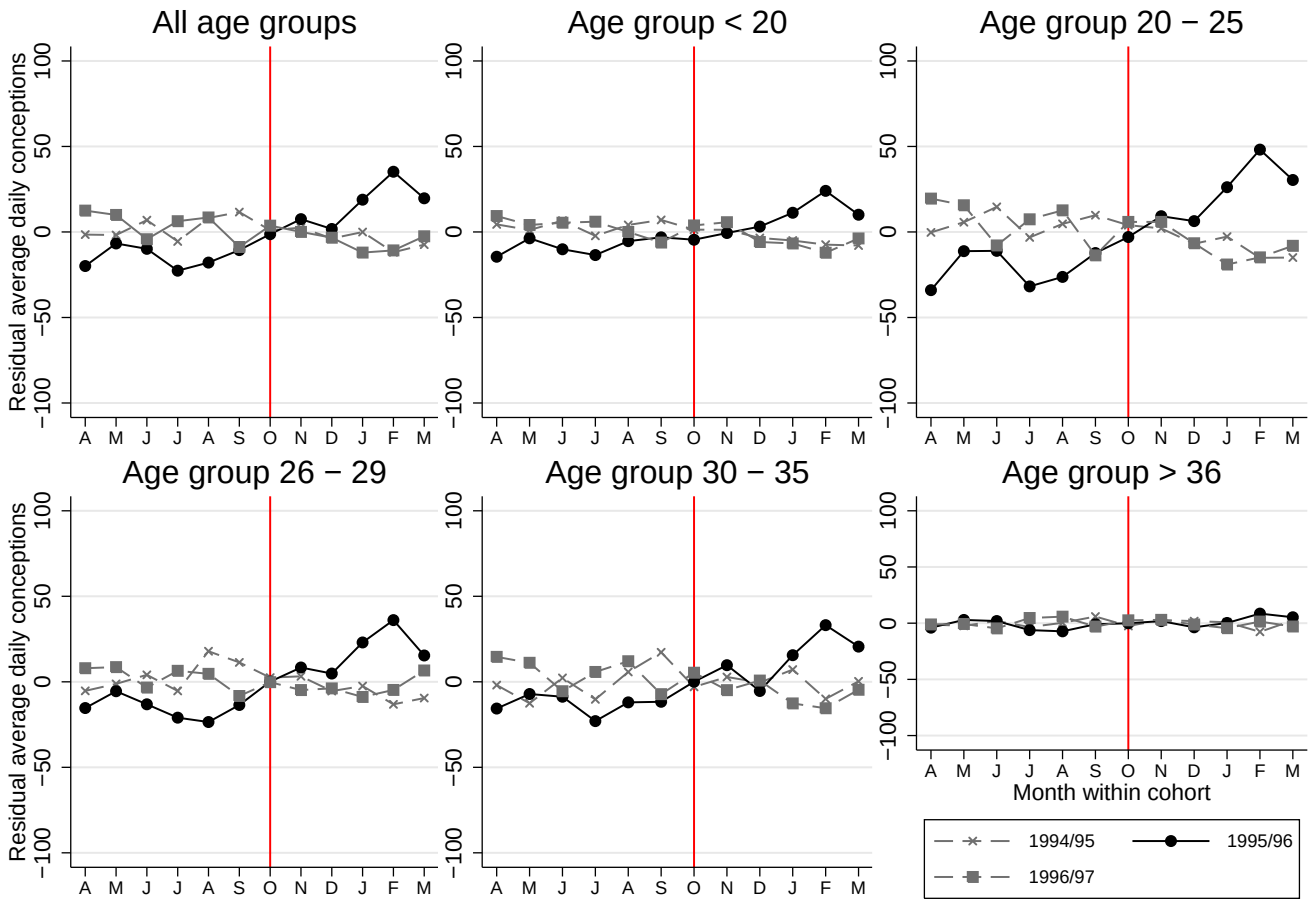
The methodological contribution illustrates how the RD approach can be a powerful method to provide causal inference in all such circumstances in which assignment to treatment depends on an observable variable and there exists a known point within its support where the probability of being treated changes discontinuously. Many other epidemiological applications could then be usefully analysed using the RD design. These include: the health effects of smoking bans, congestion charges or natural disasters (where the running variable distinguishes the areas affected by the ban, the charge, or the disaster from other control areas), the effects of parental leave policy reforms (where the selection variable is the time in which parents become eligible to receive the new leave conditions) on maternal well-being, breastfeeding and child health, and the effects of being admitted to a selective school or gain a scholarship to participate in higher education on subsequent health outcomes (where the assignment variable distinguishes pupils who enrol in selected school from control pupils who do not).

Figure 1: Age-specific daily conceptions before and after October 1995, over calendar time



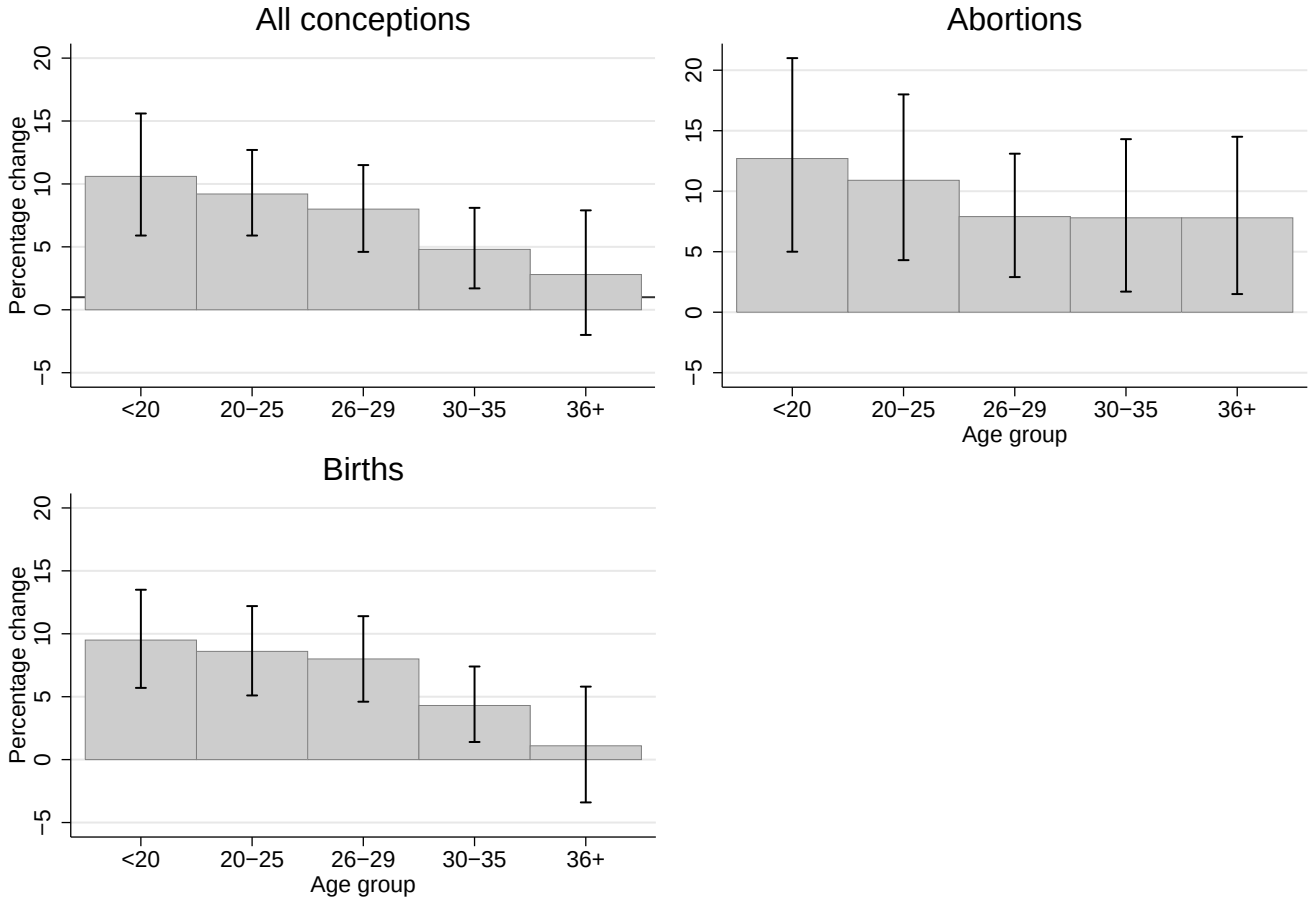
Source: ONS monthly data on conceptions (special extraction).
Notes: The data plotted are residuals from age-specific weighted Poisson regressions of daily number of conceptions, calculated by dividing monthly aggregates on number of days per month, on a full set of calendar month dummies. The weights are given by the total number of births observed for each month and mother's age group. The graph for all age groups is obtained using a weighted average of the age-specific residuals, where the weights are given by the age-specific number of births divided by the total number of births for each month.

Figure 2: Age-specific daily conceptions before and after October (month 0), by cohort



Source: ONS monthly data on conceptions (special extraction).
Notes: The data plotted are residuals from age-specific weighted Poisson regressions of daily number of conceptions, calculated by dividing monthly aggregates on number of days per month, on a full set of calendar month dummies and cohort dummies. The weights are given by the total number of births observed for each cohort, month, and mother's age group. The graph for all age groups is obtained using a weighted average of the age-specific residuals, where the weights are given by the age-specific number of births divided by the total number of births for each month and cohort.

Figure 3: Percentage change during 5 month period after October 1995 in conceptions, abortions and births by age group



Source: ONS monthly data on all conceptions, abortions and births (special extraction).
Notes: The figures show the estimates of the percentage change in all conceptions, abortions and births due to the pill scare over the May-Mar (± 5 months) time window. Point estimates are represented by the bars, while the 95% CIs are shown by the vertical lines.

Table 1: The Effect of the Pill Scare on Birth *Quantity*

Source/Outcome	Aug-Dec (±2 months)	Jul-Jan (±3 months)	Jun-Feb (±4 months)	May-Mar (±5 months)
Panel A: ONS Aggregate Data[†]				
All conceptions	5.7** [5.0–6.4]	6.2** [4.4–7.9]	7.7* [4.1–11.4]	7.4** [4.2–10.7]
Abortions	6.4 [-0.0–13.3]	8.2** [2.6–14.0]	10.8** [4.8–17.1]	9.5** [4.2–15.1]
Births	5.5** [3.4–7.7]	5.7** [3.8–7.6]	7.0** [3.6–10.6]	6.9** [3.9–10.0]
Panel B: Individual Birth Records[‡]				
Births	5.6* [3.8–7.4]	5.7** [4.0–7.4]	6.0** [4.1–8.0]	6.0** [4.0–8.1]

Sources: [†]ONS monthly data on conceptions and percentage of conceptions terminated by abortion (special extraction). [‡]Individual Birth Register data aggregated at the monthly level by age group of the mother.

Notes: Estimates are obtained from weighted Poisson regressions, where the weights are given by the total number of births observed by mother's age group, a, cohort, c, and month, m. 95% CI in square brackets. Standard errors clustered by month within cohort. Figures are percentage changes in the dependent variable. All regressions include age, cohort, age and cohort interactions, an age-specific function of month within cohort (specified as a linear function for the 2-month window, a quadratic function for the 3-month window, a cubic function for the 4-month window and a quartic function for the 5-month window), a full set of interactions between cohort dummies and an indicator variable for the period following October within each cohort. The effect of the pill scare (shown) is given by the interaction of the post-October dummy and the 1995/96 cohort dummy.

**, * denote estimate is significant at 1% and 5% level, respectively.

Table 2: The Effect of the Pill Scare on Birth *Quality*

Outcome	Aug-Dec (±2 months)	Jul-Jan (±3 months)	Jun-Feb (±4 months)	May-Mar (±5 months)
Still births	0.892 [0.755–1.053]	0.961 [0.835–1.105]	0.978 [0.879–1.089]	0.974 [0.880–1.077]
Congenital anomalies	0.857** [0.771–0.954]	0.836** [0.769–0.909]	0.889* [0.804–0.983]	0.893* [0.814–0.980]
Birth weight (grams)	6.057* [01.885–10.229]	0.731 [-10.041–11.503]	0.103 [-7.296–7.502]	4.214 [-4.774–13.203]
Very low birth weight	1.093 [0.954–1.253]	1.055 [0.956–1.164]	1.069 [0.978–1.168]	1.026 [0.935–1.127]
Low birth weight	1.027** [1.022–1.033]	1.026** [1.013–1.040]	1.021* [1.003–1.039]	0.996 [0.957–1.035]
High birth weight	1.040 [0.990–1.093]	1.035* [1.003–1.068]	1.018 [0.970–1.068]	1.019 [0.974–1.065]
Multiple births	1.119** [1.046–1.198]	1.090* [1.018–1.168]	1.056 [0.989–1.127]	1.012 [0.935–1.095]
Male child	0.997 [0.989–1.005]	1.001 [0.993–1.009]	0.994 [0.983–1.006]	0.993 [0.983–1.004]

Source: See Table 1.
Notes: Estimates are obtained from logit regressions for grouped data (and groups are given by the number of live and still births in each cohort month) for all outcomes, except birth weight, for which a weighted linear regression is performed. 95% CI in square brackets. Standard errors clustered by month within cohort. Odds ratios are reported for all outcomes, except for birth weight, for which coefficients are shown. For further information, see the note to Table 1.
**, * denote estimate is significant at 1% and 5% level, respectively.

Table 3: Robustness checks

Panel A: Months of October and November omitted, symmetric time window				
	Aug-Jan (±2 months)	Jul-Feb (±3 months)	Jun-Mar (±4 months)	
All conceptions	7.0** [4.9–9.2]	8.9** [4.6–13.5]	8.8** [5.5–12.3]	
Abortions	8.4 [-0.1–17.5]	11.3** [3.4–19.8]	11.5** [5.6–17.7]	
Births	6.6** [4.4–8.6]	8.3** [4.2–12.6]	8.2** [5.0–11.5]	
Panel B: Month of October omitted, asymmetric time window				
		Jul-Jan (±3)	Jul-Mar (−3/+ 5)	Jul-Jun (−3/+ 8)
All conceptions		6.2** [4.4–7.9]	8.3** [5.3–11.5]	7.8** [5.5–10.1]
Abortions		8.2** [2.6–14.0]	10.7** [5.0–16.7]	9.9** [4.5–15.7]
Births		5.7** [3.8–7.6]	7.7** [4.6–10.9]	7.2** [4.9–9.6]
Panel C: Differential quadratic trend before and after October, symmetric time window				
	Aug-Dec (±2 months)	Jul-Jan (±3 months)	Jun-Feb (±4 months)	May-Mar (±5 months)
All conceptions	5.7** [5.0–6.4]	6.2** [4.4–7.9]	7.7** [4.1–11.4]	7.4** [4.2–10.7]
Abortions	6.3 [-0.1–13.2]	8.2** [2.6–14.0]	10.8** [4.8–17.1]	9.5** [4.2–15.1]
Births	5.5** [3.4–7.7]	5.7** [3.8–7.6]	7.0** [3.6–10.6]	6.9** [3.9–10.0]
Panel D: Effect of dummy before/after October 1995, symmetric time window				
	Aug-Dec (±2 months)	Jul-Jan (±3 months)	Jun-Feb (±4 months)	May-Mar (±5 months)
All conceptions	3.6** [1.7–5.6]	4.9** [2.7–7.0]	5.6** [2.9–8.3]	5.7** [3.4–8.1]
Abortions	6.4** [4.7–8.2]	8.5** [5.9–11.2]	8.2** [4.2–12.4]	8.1** [5.0–11.4]
Births	3.0* [0.7–5.4]	4.1** [1.9–6.4]	5.0** [2.4–7.7]	5.2** [2.8–7.5]

Source: See Table 1.

Notes: Estimates are obtained from weighted Poisson regressions. 95% CI in square brackets. Standard errors clustered by month within cohort. Figures are percentage changes in the dependent variable. For further information on panels A and B, see note to Table 1. Panel C includes an age-specific function of month within cohort which varies before and after October of each cohort-year. Panel D ignores the subdivision in cohort-years and instead includes a dummy variable that takes value one for the period after October 1995 and zero otherwise. It also include an age-specific trend which varies before and after October 1995.

**, * denote estimate is significant at 1% and 5% level, respectively.

References

Allison, C and Roizen J. 1996. Contraception: the aftermath of the oral contraceptive 'scare'. *British J Sexual Medicine*; **23**: 13-16.

Angrist JD, Krueger AB. 1999. Empirical strategies in labor economics. In Ashenfelter O, Card D (eds.) *Hand of Labor Econ*, Vol 3. Amsterdam: North Holland.

Angrist JD, Lavy V. 1999. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly J Econ*; **114**: 533–75.

Battistin E, Brugiavini A, Rettore E, Weber, G. 2009. The retirement consumption puzzle: evidence from a Regression Discontinuity approach. *Am Econ Rev*; **99**: 2209-2226.

Campbell DT. 1969. Reforms as experiments. *Am Psych*; **24**: 409-29.

Chay KY, Greenstone M. 2005. Does air quality matter? Evidence from the housing market. *J Pol Econ*; **113**: 376-424.

Chen S, van der Klaauw W. 2008. The work disincentive effects of the disability insurance program in the 1990s. *J Econometrics*; **142**: 757–84.

Cook TD, Campbell DT. 1979. *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Boston, MA: Houghton-Miffin.

Davey-Smith, G. 2008. How do we know, what do we know and what can knowledge do? From John Brownlee to translational medicine. *Int J Epidemiol*; **37**: 911–913.

DiNardo J, Lee DS. 2004. Economic impacts of new unionization on U.S. private sector employers: 19842001. *Quarterly J Econ*; **113**: 376-424.

Farmer RDT, Williams TJ, Simpson EL, Nightingale AL. 2000. Effect of 1995 pill scare on rates of venous thromboembolism among women taking combined oral contraceptives: analysis of General Practice Research Database. *BMJ*; **321**: 477-479.

Furedi, A. 1999. The public health implications of the 1995 'pill scare'. *Human Reproduction Update*; **5**: 621–626.

Galea S, Riddle M, Kaplan GA. 2010. Causal thinking and complex system approaches in epidemiology. *Int J Epidemiol*; **39**: 97–106.

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

Hope, S. 1996. 12% of women stopped taking their pill immediately they heard CMS's warning. *BMJ*; **312**: 576.

Imbens GW, Angrist JD. 1994. Identification and estimation of local average treatment effects. *Econometrica*; **62**: 467-76.

Imbens GW, Lemieux T. 2008. Regression discontinuity designs: a guide to practice. *J Econometrics*; **142**: 615–356.

Jacobs BA, Lefgren L. 2004. Remedial education and student achievement: a regression-discontinuity analysis. *Rev Econ Stats*; **86**: 226–44.

Jick H, Jick SS, Gurewich V, Myers MW, Vasilakis C. 1995. Risk of idiopathic cardiovascular death and non-fatal venous thromboembolism in women using oral contraceptives with differing progestogen components. *The Lancet*; **346**: 1589-93.

Lee DS. 2008. Randomized experiments from non-random selection in U.S. House elections. *J Econometrics*; **142**: 675-97.

Lee DS, Card D. 2008. Regression discontinuity inference with specification error. *J Econometrics*; **142**: 655-674.

Lee DS, Lemieux T. 2010. Regression discontinuity designs in economics. *Journal of Economic Literature*; **48**: 281-355.

Loewenson, R. 2004. Epidemiology in the era of globalization: skills transfer or new skills? *Int J Epidemiol*; **33**: 1144–1150.

Maldonado G, Greenland S. 2002. Estimating causal effects. *Int J Epidemiol*; **31**: 422–29.

Pearce, N. 2004. The globalization of epidemiology: introductory remarks. *Int J Epidemiol*; **33**: 1127–1131.

Pearce, N. 2011. Epidemiology in a changing world: variation, causation and ubiquitous risk factors. *Int J Epidemiol*; forthcoming.

Poulter NR, Chang CL, Farley TMM, Meirek O, Marmot MG. 1995. World Health Organization collaborative study of cardiovascular disease and steroid hormone contraception. Venous thromboembolic disease and combined oral contraceptives: results of international multicentre case control study. *The Lancet*; **346**: 1575-82.

Ramsay S. 1996. UK "pill scare" led to abortion increase. *The Lancet*; **347**: 1109.

Robins JM, Hernán MA, Brumback B. 2000. Marginal structural models and causal inference in epidemiology. *Epidemiology*; **11**: 550–56.

Rothman KJ, Greenland S. 2005. Causation and causal inference in epidemiology. *Am J Public Health*; **95**: S144–45.

Rubin, DB. 2008. For objective causal inference, design trumps analysis. *Annals of App Stat*; **2**: 808–840.

Rubin, DB. 2006a. Causal inference through potential outcomes and principal stratification: applications to studies with 'censoring' due to death. *Statistical Science*; **21**: 299–309. (b)

Rubin, DB. 2006b. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Statistics in Medicine*; **26**: 20-30. (a)

Spitzer WO, Lewis MA, Heinemann LA, Thorogood M, MacRae KD. 1995. Third generation oral contraceptives and risk of venous thromboembolic disorders: an international case-control

study. Transnational Research Group on Oral Contraceptives and the Health of Young Women. *BMJ*; **312**: 83-88.

Stross, JK, Harlan, WR. 1979. The dissemination of new medical information. *JAMA*; **241**: 2622–2624.

Susser M. 1991. What is a cause and how do we know one? A grammar for pragmatic epidemiology. *Am J Epidemiol* **133**: 635–48.

Thislethwaite D, Campbell D.1960. Regression-discontinuity analysis: an alternative to ex post facto experiment. *J Educ Psych*; **51**: 309–17.

Thomas, J. 2010. Risk and panics: national newspaper coverage of the cases of the contraceptive pill, drug facilitated sexual assault, dangerous dogs and road rage in the United Kingdom. Unpublished D.Phil thesis, Department of Sociology, University of Oxford.

Trochim WK. 1984. *Research Design for Program Evaluation: The Regression-Discontinuity Approach.* Beverly Hills, CA: Sage.

van der Klaauw W. 2002. Estimating the effect of financial aid offers on college enrolment: a regression-discontinuity approach. *Int Econ Rev*; **43**: 1249–87.

van der Klaauw, W. 2008. Regression-discontinuity analysis: a survey of recent developments in economics. *Labour*; **22**: 219–45.

Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. 2002. Segmented regression analysis of interrupted time series studies in medication use research. *J of Clil Phar and Therap*; **27**: 299-309.

Wood R, Botting B, Dunnell K. 1997.Trends in conceptions before and after the 1995 pill scare. *Pop Trends*; **89**: 5-12.