



Occasional Paper Series

Number: 2005-15

Title: **Matching Estimates of the Impact of Over-the-Counter
Emergency Birth Control on Teenage Pregnancy**

Authors: Sourafel Girma and David Paton

Abstract: In this paper we demonstrate how matching estimators can be used to evaluate policy interventions which are implemented in relatively few regions at different times. Our technique is based on translating calendar time into 'experimental time' to provide a common starting point for entry by different areas into the scheme. Such an approach is likely to have many applications, in particular to cases of state- or country-level interventions for which only aggregate data are available. We illustrate the technique using the case of free over-the-counter access to emergency birth control (EBC) for teenagers in England. We construct matching estimates of the impact of this scheme on the under-18 conception rate in local authorities. Irrespective of either the matching or the adjustment procedure, we do not find evidence that pharmacy EBC schemes led to significantly lower teenage pregnancy rates.

Keywords: matching estimators; family planning; teenage pregnancy; emergency birth control.

JEL-Classification: C21, I18, J13

Publication Status: unpublished

WWW: <http://www.nottingham.ac.uk/~lizecon/RePEc/pdf/matching.pdf>

Matching Estimates of the Impact of Over-the-Counter Emergency Birth Control on Teenage Pregnancy

NUBS Working Paper Series

October 2005

Sourafel Girma

Nottingham University Business School
Jubilee Campus
Wollaton Road
Nottingham
NG8 1BB
United Kingdom

Tel: + 00 44 115 846 6656

Fax: + 00 44 115 846 6667

Email: Sourafel.Girma@nottingham.ac.uk

and

David Paton*

Nottingham University Business School
Jubilee Campus
Wollaton Road
Nottingham
NG8 1BB
United Kingdom

Tel: + 00 44 115 846 6601

Fax: + 00 44 115 846 6667

Email: David.Paton@nottingham.ac.uk

*Corresponding author

Acknowledgements: We would like to thank attendants at the 14th European Workshop on Econometrics and Health Economics in Dublin and, in particular, the discussant, David Bradford for providing us with many useful insights and comments. Thanks are also due to Angela Edwards at the Teenage Pregnancy Unit, Ian Saberton at the Department of Health and other staff at the Office of National Statistics for help with data, as well as various Teenage Pregnancy Coordinators and Health Authority staff who provided information and/or comments.

Abstract

In this paper we demonstrate how matching estimators can be used to evaluate policy interventions which are implemented in relatively few regions at different times. Our technique is based on translating calendar time into ‘experimental time’ to provide a common starting point for entry by different areas into the scheme. Such an approach is likely to have many applications, in particular to cases of state- or country-level interventions for which only aggregate data are available. We illustrate the technique using the case of free over-the-counter access to emergency birth control (EBC) for teenagers in England. We construct matching estimates of the impact of this scheme on the under-18 conception rate in local authorities. Irrespective of either the matching or the adjustment procedure, we do not find evidence that pharmacy EBC schemes led to significantly lower teenage pregnancy rates.

Keywords: matching estimators; family planning; teenage pregnancy; emergency birth control.

JEL Classifications: C21, I18, J13.

Ethics: as this study deals with an analysis of publicly available data at an aggregate level, there are no relevant ethical issues.

Original Publication: we confirm that this manuscript contains original unpublished work and is not being submitted for publication elsewhere at the same time.

Matching Estimates of the Impact of Over-the-Counter Emergency Birth Control on Teenage Pregnancy

I. Introduction

There is a large and growing literature on the use of matching estimators to evaluate the impact of a policy change. The basic motivation for such studies is that difference-in-difference estimates of policy change are likely to be biased if the control group differs in systematic ways from the intervention group. In general these estimators are applied to cases in which a policy change affects a large number of subjects (the intervention group) at the same time, whilst leaving other subjects (the control group) unaffected. In many cases, most particularly those based on state-level data, policy changes are introduced by different states at different times and the number of policy changes in any one time period is insufficient to use matching estimators in a meaningful way. An obvious example which has been the subject of much recent investigation is state-level restrictions on access to abortion such as parental notification laws (see Levine, 2004, for an overview of this literature).

Policy evaluations in such cases have tended to rely on panel data difference-in-difference estimators without an explicit treatment of the matching problem. In this paper we demonstrate how matching estimators can be applied to a situation where a policy change is introduced in different areas at different times. The case we use to examine the technique is the introduction of schemes to provide emergency birth control (EBC) free of charge over-the-counter to teenagers via pharmacies in England. Our technique is based on translating calendar time into ‘experimental time’ to provide a common starting point for entry by different areas into the scheme. We then use standard matching estimators to evaluate whether areas that introduced the policy experienced significant decreases in rates of teenage pregnancy relative to areas that did not introduce the policy.

The rest of the paper is set up as follows. In section II we introduce the policy background for over-the-counter emergency birth control and briefly review existing

evaluations of the policy. In section III, we explain our matching technique in more detail and also the data used. We present the results in section IV and make some concluding comments in section V.

II. Over-the-counter Emergency Birth Control

There is a long-running debate over whether emergency birth control, sometimes known as the morning after pill or emergency contraception, should be made available over the counter from pharmacists and other sources. Boonstra (2002) claims that wider access to emergency birth control contraception “is perhaps the single most promising avenue for reducing this country's high rates of unintended pregnancy and abortion” and is “especially important for sexually active teens” (p.13). On the other hand, theoretical models of decision making suggest that greater access to family planning services may increase risk taking behaviour and that the overall impact on unwanted pregnancy rates is impossible to ascertain *a priori* (see, for example, Akerlof et al, 1996; Paton, 2002).

In contrast to the situation in the USA where the FDA has refused to allow general over-the-counter access to EBC, many European countries have legislated to permit pharmacies and other venues to provide EBC without a prescription. Since 2000, a key element in the English strategy for reducing teenage pregnancy has been to improve access to emergency birth control (EBC) for young people. Starting in early 2000, local areas have been permitted to fund schemes in which pharmacists provide EBC over-the-counter free of charge to teenagers. There is no requirement for parental consent and in most cases there is no specific lower age limit for accessing the service.

We report the development of participation in the scheme in Table 1. By the end of 2002, 58 (nearly 40%) of all local authorities in England were participating in the pharmacy EBC scheme for teenagers. Unfortunately, there is no systematic collection of data on the

take-up of EBC from these schemes. However, a few local authorities make available at least some information on the take-up and this helps to give an idea of their potential impact. A typical example is the London borough of Enfield. Between 1998 and 2001, Enfield experienced an average of 217 under-18 conceptions per year. This equates to a conception rate of 44.9 per thousand women aged 15 to 17, fairly close to the mean for the whole of England. Enfield entered the pharmacy EBC scheme in 2002 and the local teenage pregnancy co-ordinator produces a detailed breakdown of take-up by age. Each year about 300 under-18s are provided with EBC free of charge under the pharmacy scheme, a rate of about 62 per thousand. Using the mid-point of the EBC effectiveness estimates provided by Trussell et al (1998), we would expect about 12 pregnancies to have been avoided by this level of use, implying a reduction in the under-18 pregnancy rates of just over 5%, fully one third of the target reduction in the under-18 rate for Enfield by 2004 of 15%.

To date there have been very few evaluations of over-the-counter access to EBC. Indeed, the existing evidence base on the impact of access to EBC on unwanted pregnancy is limited to just four papers. Two of these papers report the results from randomised control trials (RCTs) based on groups recruited from attendees at family planning clinics. Glasier and Baird (1998) report an RCT on the effect of self-administering EBC amongst women of all ages. Although the rate of unwanted pregnancy decreased relative to the control group, the difference was not statistically significant. Raine et al (2005) examine the impact of access to EBC through pharmacies and via advance provision on pregnancy rates of women under the age of 25. The authors found that, compared with the control group, women in the pharmacy access and advance provision groups did not experience a significant reduction in pregnancy rate. This appears to be the only RCT to date which has examined the impact of emergency birth control on pregnancy rates among young women.

Unfortunately, it is difficult to extrapolate from these results, the likely impact of easier access to EBC on unwanted pregnancy amongst the population as a whole. In the first place, pregnancy is a relatively rare event and it is difficult to achieve a large enough sample size to ensure that the subsequent statistical tests are of high power. Further, there is an inherent sample selection bias. Specifically, existing users of family planning services (the population from which the samples in these papers were taken) may exhibit a different response to changes in EBC access to other sexually active young people. Further, restricting subjects to those who are already sexually active makes it impossible to observe any impact of the policy change on other young people. This latter point implies that estimates of pregnancy reduction from access to EBC in these papers represent upper bounds at best.

Two papers have examined the impact of pharmacy access to EBC on pregnancy rates at the population level. Glasier et al (2004) compare changes in abortion and birth rates between 1998–1999 and 2000-2001 amongst young women in one area of Scotland in which advance provision of emergency birth control was introduced and in other areas with no such provision. The authors find no significant differences in changes in birth rates or abortion rates between the areas before and after the scheme. Paton (2005) uses a 2-way fixed effects panel specification to examine the impact of a range of family planning measures on teenage conception rates in 99 local authorities in England between 1998 and 2001, one such measure being the presence of schemes to provide free EBC to young people. In general, pharmacy EBC was not associated with any significant change in pregnancy rates, although significant increases in diagnoses of sexually transmitted infections (STIs) amongst young people were detected in some specifications. The results of this study are somewhat limited by the fact that they cover only the first two years of operation of the pharmacy-EBC scheme, in which relatively few areas had entered the scheme.

Although the two population studies employ a difference-in-difference methodology in that changes over time in intervention areas are compared to changes over time in non-intervention areas, in neither study was there any attempt to ensure that the control and intervention groups were matched in any way. There is an extensive literature which discusses the difficulty of identifying the impact of policy changes using regional and time variations (see Besley and Case, 2000 for a general discussion). Fundamentally, the problem arises if policy changes are influenced either by the outcome variable or by control variables. In our context, the problem is that heterogeneity between areas choosing to operate a pharmacy EBC scheme and others may be associated with pregnancy rates. For example, areas with high rates of teenage pregnancy will be more likely to implement measures to reduce these rates. To some extent, this problem is alleviated by the inclusion of county fixed effects which can control for time-invariant fundamental differences in county pregnancy rates. However, the endogeneity problem may be more subtle. Areas with high teenage pregnancy rates may contain teenagers whose behaviour is relatively difficult to change. If so, the impact of pregnancy-reduction measures will have least effect in these areas and this may lead to a bias in the estimate of the policy impact. In the next section we discuss how this problem might be tackled in the context of the English pharmacy EBC scheme.

III. Empirical Methodology

Let $PHM_{it} \in \{0,1\}$ be an indicator of whether a local authority (LA) i has participated in the pharmacy EBC scheme at time period t , and let y_{it+s}^1 be the outcome variable (e.g. pregnancy rates) at time $t+s, s \geq 0$ following programme participation. Also let y_{it+s}^0 be the pregnancy rate (or any other outcome variable of interest) had the LA not participated in the scheme. The causal effect of the scheme for LA i on the outcome variable at time period $t + s$ is then defined as:

$$y_{it+s}^1 - y_{it+s}^0. \quad (1)$$

The fundamental problem of causal inference is that the quantity y_{it+s}^0 which is the counterfactual is inherently unobservable. Thus to make the problem tractable, we concentrate on identifying the average effect of the scheme on participants. This effect is known in the microeconomic evaluation literature as *the average effect of the treatment on the treated* (e.g. Heckman et al, 1997, Dehejia and Wahba, 2002 and Smith and Todd, 2005a). Mathematically, the average effect of the treatment on the treated is defined as

$$E \left\{ y_{t+s}^1 - y_{t+s}^0 \mid PHM_{it} = 1 \right\} = E \left\{ y_{t+s}^1 \mid PHM_{it} = 1 \right\} - E \left\{ y_{t+s}^0 \mid PHM_{it} = 1 \right\}. \quad (2)$$

We refer to the resulting estimator as the *standard matching estimator*.

Causal inference relies on the construction of the counterfactual for the last term in equation (2), which is the outcome scheme participants would have experienced, on average, had they not been involved in the scheme. This is estimated by the corresponding average value of the outcome variable for the LAs that did not participate in the scheme,

$$E \left\{ y_{it+s}^0 \mid PHM_{it} = 0 \right\}.$$

When there are repeated observations for the same set of participants over time, it was found more reliable to base the policy evaluation analysis on the difference between the variable of interest (viz. y_{it+s}^1) and its pre-treatment value (viz. y_{it-1}^1), that is

$\Delta y_{it+s}^1 = y_{it+s}^1 - y_{it-1}^1$ (e.g. Blundell and Costa Dias, 2000). In this case the *average effect of the treatment on the treated* is defined as

$$E \left\{ \Delta y_{t+s}^1 - \Delta y_{t+s}^0 \mid PHM_{it} = 1 \right\} = E \left\{ \Delta y_{t+s}^1 \mid PHM_{it} = 1 \right\} - E \left\{ \Delta y_{t+s}^0 \mid PHM_{it} = 1 \right\} \quad (3)$$

Since the resulting estimator is based on differences, it is known as the *difference-in-differences matching estimator*.

An important feature in the accurate construction of the counterfactual is the selection of a valid control group. The approach we take here is to employ matching techniques. The

purpose of matching is to pair each participating LA with a non-participating one on the basis of some observable variables, in such a way that the outcome variable of the latter can be studied to generate the counterfactual for the former. This type of matching procedure is preferable to randomly or indiscriminately choosing the comparison group, because it is less likely to induce estimation bias by picking LAs with markedly different characteristics. In practice matching involves comparing participating and non-participating LAs across a number of observable pre-participation characteristics, in particular levels and trends in teenage pregnancy, ethnic mix, existing provision of sexual health services and so on. Thus it is desirable to perform the matching on the basis of a single index that captures all the information in those variables. To this end, we adopt the method of propensity score matching due to Rosenbaum and Rubin (1983). This involves the use of the probability of participating in the scheme conditional on those characteristics, to reduce the dimensionality problem. Accordingly, we first identify the probability (or *propensity score*) of programme participation using a probit model

$$P(PHM_{it} = 1) = F(X_{it-1}) \quad (4)$$

where X is a vector of covariates observed in the time period before programme participation,

Now let p_i denote the predicted probability of programme participation for LA i amongst programme participants (say group A) and let p_j denote the predicted probability of programme participation for firm j in the control group of LAs (say group C). In general the difference-in-differences matching estimator of the causal effect of programme participation can be written as

$$\mu = \sum_{i \in A} \left(\Delta y_i - \sum_{j \in C} g(p_i, p_j) \Delta y_i \right) \quad (5)$$

where $g(\cdot)$ is a function assigning the weights to be placed on the comparison firm j used as a match for programme participant i . The different matching estimators proposed in the

literature (such as the nearest neighbours and kernel estimators) differ from each other in the choice of the weighting function they employ. However, they share the same property of being consistent estimators of the treatment effect under consideration.

The theoretical properties of the various average treatment effect estimators are developed in the context where all subjects receive the treatment at the same point in time. If there are several new programme participants at each point in time it is preferable to conduct a period by period analysis. In common with many contexts in which the units of analysis are regions (as opposed to individuals), such an approach is not feasible here due to the small numbers of regions entering the scheme at any one time. Consequently, we pool the available information around some ‘experimental’ time frame. For example $t = 0$ would refer to the time period the LA participated in the programme (irrespective of calendar time) and $t = -1$ would denote the pre-treatment period. We then use the difference-in-differences matching estimator described in Equation (5) to estimate the contemporaneous as well as lagged programme effects. As we have quarterly pregnancy data available to us, t indicates quarters and we estimate programme effects for up to 8 consecutive quarters.

This raises the problem of how we should measure differences for the LAs who do not participate in the programme. We do this by randomly allocating $t = 0$ to a real time period. We consider two methods of doing this. The first is a straight random allocation. The second is a random allocation, stratified to ensure the same ‘real time’ distribution of LAs in the intervention and control groups. For example, if 10% of LAs who eventually participate first enter the scheme in quarter 1 of 2001, then we stratify the sample to ensure that for 10% of the control group, $t = 0$ is allocated to quarter 1 of 2001. In practice we find the results are very similar whichever allocation method we use and the results below are based on the straight random allocation. We also performed the allocation process several times to confirm that the results are not sensitive to its outcome.

Although matching estimators are consistent, they are found to exhibit finite-sample biases. Accordingly, we also implement the small-sample biased-adjusted matching estimator due to Abadie and Imbens (2002) by way of robustness check. The bias-corrected matching estimator adjusts the difference within the pairs of matched LAs for the differences in the values of other factors affecting the outcome variables. In the present context this is done by adjusting the outcome variable (say pregnancy rate at time t) to differences in a number of covariates (described below). Such an adjustment is done by first running a non-parametric regression of the outcome variable on the covariates and then taking the residual. For this reason the biased-adjusted estimator is also known as regression-adjusted estimator.

An important issue to consider is the impact of spatial correlation. In particular, if neighbouring authorities already operate schemes, two effects may operate due to the fact that EBC schemes normally do not operate any restriction on the place of residence of clients. Firstly, a local authority may have an incentive to ‘free ride’ and be less likely to invest scarce resources in an EBC scheme when close neighbours already operate such schemes. Secondly, some teenagers resident in a local authority classified as being a control due to the absence of an EBC scheme may in fact be ‘treated’ by availing themselves of schemes in neighbouring areas.

It is possible that this problem is not very great in practice as in most cases, the costs of travel between local authorities will be non-negligible for teenagers under the age of eighteen and who are unlikely to have access to independent means of transport. A likely exception is the London region within which the population tends to be highly mobile (e.g. it is common for a school to be in a different authority to the place of residence) and the public transport network between local authorities is extensive and well-used. For this reason, we perform our analysis both with and without local authorities within the London region. A

further check that we perform is to include a variable in our regression adjusted estimates that controls for the existence of EBC schemes in neighbouring local authorities.¹

Testing the reliability of the propensity score matching method

The propensity score matching method will provide a reliable and robust method for estimating programme effects if, conditional on the propensity score, the distribution of the pre-programme covariates is independent of the incidence of programme participation. This can be achieved by choosing a specification of the propensity score model (cf. Equation 4) that ‘balances’ the pre-programme variables between the treatment and control groups conditional on the propensity score. As emphasised by Rosenbaum and Rubin (1983) and Dehejia and Wahba (2002) it is important to verify that this balancing condition is satisfied by the data. In this paper we perform three balancing tests suggested in the recent literature (e.g. Smith and Todd, 2005b).

The first balancing test examines the standardised difference (or bias) for all variables in X (that is the vector of covariates used in the propensity score estimation) as described in Smith and Todd (2005b). For example, the standardised bias for the *ethnic* variable is defined as the difference in means between programme participants (group A) and the matched comparison group of LAs (group C), scaled by the average variances of the *ethnic* variable for groups A and C. Based on N programme participants, this is given as

$$SDIFF(ethnic) = \frac{100 \frac{1}{N} \sum_{i \in A} \left[ethnic_i - \sum_{j \in C} g(p_i, p_j) ethnic_j \right]}{\sqrt{\frac{Var_{i \in A}(ethnic) + Var_{j \in C}(ethnic)}{2}}}. \quad (6)$$

Note that the lower the standardised difference, the more balanced or similar the treatment and comparison groups will be in term of the variable under consideration.

¹ Note that including such a variable in the vector used to match the programme and non-programme

Although a formal criterion as to how large a standardised bias should be for it to be considered serious does not exist, we follow Rosenbaum and Rubin (1985) and assume that a value of 20 is large. Furthermore, for each variable entering the propensity score model we perform standard t-tests of equality between treated and control firms to satisfy ourselves that no significant differences exist.

Whereas the above balancing test considers the cross-sample difference of each variable entering the probit model separately, there also exists a test that considers whether those differences can be taken as *jointly* insignificant. This test is known as the Hotelling's T-squared test, and it has the flexibility of being based either on all observations or for separate segments of the sample defined by the propensity score estimates. In this study we divide the sample into four equal parts (i.e. by propensity score quartile), and conduct the Hotelling's T-squared test within each part.

The third balancing test we explore is suggested by Todd and Smith (2005b) and it is cast within a regression framework. Let $\hat{P}(X)$ denote the estimated propensity score and let D be a dummy variable assuming a value of 1 if a LA is programme participant. Then for each variable included in the matching algorithm, the following regression function that is quartic in $\hat{P}(X)$ is estimated (again using the *ethnic* variable as an example):

$$Ethnic = \beta_0 + \sum_{k=1}^4 \beta_k \hat{P}(X)^k + \sum_{k=1}^4 \gamma_k D \hat{P}(X)^k + \varepsilon \quad (7)$$

and the joint significance of the coefficients on the terms involving the programme participation dummy (that is the γ s) is tested. As explained by Todd and Smith (2005b), if the propensity score satisfies the balancing condition, D should not provide any additional information and we should expect the γ s to be jointly statistically insignificant.

participants could exacerbate the problem.

Throughout we impose the so-called common support condition in the matching algorithm. This involves dropping program participating LAs whose propensity score is higher than the maximum or less than the minimum propensity score of the control group of LAs.

IV. Data and Results

Data

The primary outcome measure of interest is the under-18 conception rate. Reductions in this measure are the key target in the English Teenage Pregnancy Strategy published in 1999. The under-18 conception rate is available for all 147 higher tier local authorities for each quarter between 1998 and 2002. Given that local authorities can (and do) decide to enter into the pharmacy EBC scheme at any point during the calendar year, the use of quarterly data is particularly attractive. The date of entry into the EBC scheme was confirmed by representatives of local authorities and/or health authorities within each area, the information being cross-checked with information supplied by the Teenage Pregnancy Unit.

The vector (X) used to match the programme and non-programme participants consists of the following variables, all observed in the time period before programme participation. Firstly we include the calendar year corresponding to experimental time 0 (*year*). This controls for the fact that average conception rates vary from year to year. Secondly we include the 2004 and 2010 targets reductions in teenage pregnancy set for each LA by the Teenage Pregnancy Unit (*target2004* and *target2010*). Targets are larger for local authorities with relatively high rates of teenage pregnancy. Further, the level funding provided by central Government is directly related to the target. Hence, we would expect LAs with higher targets to be more likely to enter into the pharmacy-EBC scheme than others. Next we allow for differential trends in pregnancy rates by including the change in the conception rates since 1995 (*cur18ch*). Again, we would expect areas with rapidly

increasing rates to be more likely to participate. Thirdly we include a series of variables describing the existing extent of sexual health services for young people. These are the number of clinic youth family planning sessions provided per person (*clinicr*), the number of GP practices per person (*praticer*) and the number of GP practices offering contraception services to any person (*contalla*), the proportion of teenagers attending family clinics who are provided with EBC (*EBCP*) and the proportion of young people covered by the APAUSE sex education programme (*apause*). Lastly we include two variables relating to population. Firstly the total female population aged 15-17 (*population*) and the percentage of the population in the LA who are from an ethnic minority as reported by the 2000 Census (*ethnic*).

We use the following variables in small-sample biased-adjusted matching estimates. Firstly to control for seasonal effects and for national trends in pregnancy rates, we include the year & quarter corresponding to the experimental time (*date*). Secondly, we include three socio-economic variables for the corresponding time period: female quarterly unemployment rate, annual proportion of children aged 15-17 in local authority care (*care18*) and the percentage of school pupils with no qualifications at age 16, measured as a three-year rolling average (*noqualav*). Thirdly, to allow for the possibility that neighbouring local authorities may be running schemes and that cross-border travel may be a possibility for some teens, we include a variable indicating the percentage of contiguous local authorities which have an EBC scheme in operation (*contig*). We also include several variables described earlier in the matching procedure, namely *map*, *praticer*, *contallr*, *clinicr* and *apause*.

Results

In Tables 2 and 3 we report the balancing test results based on the nearest neighbour matching method. In contrast to the unmatched sample, the standardised differences between

the treatment and comparison sample are quite moderate and statistically insignificant. The results from the Hotelling's T-squared test also provide support for the soundness of our propensity score specification.

[Tables 2 and 3]

The regression-based tests also corroborate the success of the propensity score-matching approach as we fail to reject the hypothesis that the γ s in Equation 7 are jointly statistically insignificant. It is thus comforting that both standardised differences and regression tests perform well, suggesting that the propensity score specification we have chosen is effective in accounting for factors that determine selection into treatment (i.e. programme participation).

We present the estimates of the programme effects following the basic matching process in Table 4. We present the estimated programme effects using the small-sample adjusted matching estimator in Table 5. In both cases we present results based on nearest neighbour matching and nearest three neighbours matching. At the top of each table we report the average treatment effect for the first and second year of intervention. The sign of the estimated Year 1 effect varies with the specification. However, in every case, the effect is small and statistically insignificant. The Year 2 effect is always estimated to be positive, somewhat larger in magnitude, but still insignificant.

Below, the year effects, we report the effects for each of the first eight quarters. This allows us to examine whether there is any trend in treatment effect over time. In fact, the estimated effects for the first seven quarters have inconsistent signs, are small in magnitude and statistically insignificant, at least at the 5% level.

[Tables 4 and 5]

As a further experiment, we re-estimate our models excluding all local authorities within London. As discussed above, we see this as a useful control as the effect of travel

across local authority boundaries is likely to be felt particularly in the London region. The pattern of results (not reported here) is very similar to those reported earlier with no significant (at the 5% level) negative effects estimated for any specification.

V. Discussion and Conclusions

In this paper we demonstrate how matching estimators can be used to evaluate policy interventions which are implemented in relatively few regions at different times. Such an approach is likely to have many applications, in particular to cases of state- or country-level interventions for which only aggregate data are available.

We illustrate the technique using the case of free over-the-counter access to emergency birth control (EBC) for teenagers in England. We construct matching estimates of the impact of this scheme on the under-18 conception rate in local authorities. Irrespective of either the matching or the adjustment procedure, we find little evidence that pharmacy EBC schemes have led to lower under-18 pregnancy rates in England.

Thus the results in this paper do not provide evidence that schemes allowing emergency birth control to be provided in pharmacies free of charge to young people lead to significant changes in teenage pregnancy rates. This result is consistent with previous studies of the impact of emergency birth control, including those based on randomised controlled experiments. Given the hope that many policy makers and health professionals have held out for the potential of EBC in reducing unwanted pregnancies, this finding will be disappointing.

Why should access to EBC appear to have so little impact on unwanted pregnancies? At least three possible explanations suggest themselves. The first is that pharmacy-EBC schemes simply switch the take-up of EBC away from other sources, such as family planning clinics and to pharmacies. The second possible explanation is that pharmacy-EBC schemes

do in fact result in reductions in teenage pregnancy, but that the effects are too small for the statistical tests used to reveal them, in spite of the numerous robustness checks that are conducted. The third possible explanation is that, as predicted by standard microeconomic theory, the introduction of the pharmacy-EBC schemes induced at least some adolescents to increase their level of risk-taking sexual behaviour and that the reduction in pregnancies from greater use of EBC is being countered by additional pregnancies resulting from this behaviour change.

Although a full treatment of these three possibilities is outside the scope of this present paper, a potentially fruitful possibility for future work would be to examine the differential impact of EBC schemes on teenage pregnancy and other outcomes of risky sexual behaviour, in particular, sexually transmitted infections.

REFERENCES

- Abadie, A and G Imbens (2002), *Simple and Bias-Corrected Matching Estimators*, Technical Department of Economics, UC Berkeley, <http://emlab.berkeley.edu/users/imbens/>.
- Akerlof, George, A, Yellen, Janet, L and Katz, Michael, L (1996), 'An analysis of out-of-wedlock childbearing in the United States', *Quarterly Journal of Economics*, 111 (2, May): 277-317.
- Besley, Timothy and Case, Anne (2000), 'Unnatural Experiments? Estimating the Incidence of Endogenous Policies', *The Economic Journal*, Nov., 110: F672-F694.
- Blundell, R and M Costa Dias, M (2000), 'Evaluation methods for non-experimental data', *Fiscal Studies*, 21: 427-468.
- Boonstra, Heather (2002), 'Emergency contraception: steps being taken to improve access', *The Guttmacher Report on Public Policy*, 5 (5, December): 10-13.
- Dehejia, R. and S. Wahba (2002), 'Propensity score matching methods for non-experimental causal studies', *Review of Economics and Statistics*, 84: 151-161.

- Glasier Anna and Douglas Baird (1998), 'The effects of self-administering emergency contraception', *New England Journal of Medicine*, 339: 1-4.
- Glasier, Anna, Karen Fairhurst, Sally Wyke, Sue Zirebland, Peter Seaman, Jeremy Walker and Fatim Lakha (2004), 'Advanced provision of emergency contraception does not reduce abortion rates', *Contraception*, 69: 361-6.
- Heckman, J, H Ichimura, J Smith and P Todd, (1997), 'Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme', *Review of Economic Studies*, 64: 605-654.
- Levine, Phillip B (2004), *Sex and Consequences: abortion, public policy and the economics of fertility*, Princeton: Princeton University Press.
- Paton, David (2002), 'The economics of abortion, family planning and underage conceptions', *Journal of Health Economics*, 21 (2, March): 27-45
- Paton, David (2006), 'Random behaviour or rational choice? Family planning, teenage pregnancy and STIs', *Sex Education*, forthcoming.
- Raine TR, Harper CC, Rocca CH, et al. (2005), 'Direct access to emergency contraception through pharmacies and effect on unintended pregnancy and STIs: a randomized controlled trial', *JAMA*, 293: 54-62.
- Rosenbaum, P and DB Rubin (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika*, 70: 41-55.
- Rosenbaum, P and DB Rubin (1985), 'Constructing a control group using a multivariate matched sampling method that incorporates the propensity score', *The American Statistician*, 39: 33-38
- Smith, J and P Todd (2005a), 'Does matching overcome Ladonde's critique of nonexperimental estimators?', *Journal of Econometrics*, 125: 305-353.
- Smith, J and P Todd (2005b), 'Rejoinder', *Journal of Econometrics*, 125: 365-375.

Trussell J, G Rodriguez, C Ellertson (1998), 'New estimates of the effectiveness of the Yuzpe regimen of emergency contraception', *Contraception*, 57 (6, June): 363-9

Table 1: Pattern of Entry to Pharmacy EBC Scheme in England: 2000 to 2002

Year & Quarter	New Entrants	Total LAs in Scheme	% LAs in Scheme
2000, q1	3	3	2.04
2000, q2	0	3	2.04
2000, q3	7	10	6.80
2000, q4	3	13	8.84
2001, q1	6	19	12.93
2001, q2	8	27	18.37
2001, q3	3	30	20.41
2001, q4	8	38	25.85
2002, q1	8	46	31.29
2002, q2	3	49	33.33
2002, q3	7	56	38.10
2002, q4	3	59	40.14

Table 2: Balancing tests based on nearest neighbour matching

Variable	Sample	MEAN		% bias	% bias reduction	t- test for equality of means	
		Treated	Control			t	p> t
year	Unmatched	2000.9	2000.8	7.9		0.46	0.65
	Matched	2000.9	2001.1	-24.5	-209.8	-1.2	0.235
target2004	Unmatched	14.407	14.726	-22.8		-1.38	0.17
	Matched	14.63	14.769	-10	56.4	-1.27	0.207
target2010	Unmatched	50.932	49.695	29		1.68	0.096
	Matched	50.833	49.352	34.7	-19.8	1.93	0.057
cur18ch	Unmatched	3.553	3.1775	2.4		0.14	0.885
	Matched	3.3089	5.3797	-13.3	-451.5	-0.57	0.568
contallr	Unmatched	0.02774	0.02665	16.3		0.96	0.34
	Matched	0.02721	0.02716	0.8	95.2	0.43	0.669
practicer	Unmatched	11.185	10.336	22.4		1.35	0.178
	Matched	10.801	10.572	6	73.1	0.75	0.452
clinier	Unmatched	0.03732	0.04702	-26.6		-1.5	0.135
	Matched	0.03798	0.02705	30	-12.8	1.88	0.064
map	Unmatched	0.31914	0.33875	-16		-0.96	0.34
	Matched	0.32121	0.31597	4.3	73.3	0.13	0.898
apause	Unmatched	0.00502	0.00585	-2.5		-0.14	0.887
	Matched	0.00548	0.00939	-11.7	-368	-0.8	0.424
population	Unmatched	7111.9	5704.8	30.6		1.81	0.072
	Matched	6787.4	7792.8	-21.8	28.6	-0.56	0.578
ethnic	Unmatched	10.985	11.189	-1.6		-0.09	0.927
	Matched	10.844	11.241	-3.1	-94.9	-0.1	0.924

Notes: The regression-based tests also confirm the validity of the propensity-score matching approach.

Table 3: Hotelling's T-squared tests

Quartile	T-squared statistics	F-test statistics	p-value
1	27.108	1.084	0.482
2	4.934	0.299	0.978
3	29.417	1.952	0.077

Table 4: Matching Estimates of Pharmacy EBC on U18 Pregnancy Rates

Time Period	Nearest Neighbour		Nearest 3 Neighbours	
	Effect	t-stat	Effect	t-stat
Year 1	0.167	0.111	-1.209	-0.845
Year 2	2.928	1.512	1.870	0.732
Quarter 1	-0.322	-0.590	-0.474	-0.925
Quarter 2	-0.159	-0.291	-0.003	-0.005
Quarter 3	-0.496	-0.590	-0.673	-1.207
Quarter 4	-0.080	-0.115	-0.075	-0.101
Quarter 5	0.171	0.277	0.535	0.893
Quarter 6	0.567	0.862	0.200	0.250
Quarter 7	-0.079	-0.101	-0.617	-1.026
Quarter 8	0.777	0.852	0.839	1.183

Table 5: Regression Adjusted Matching Estimates of Pharmacy EBC on U18 Pregnancy Rates

Time Period	Nearest Neighbour		Nearest 3 Neighbours	
	Effect	t-stat	Effect	t-stat
Year 1	0.327	0.156	-0.207	-0.142
Year 2	3.536	1.459	2.971	1.441
Quarter 1	-0.324	-0.722	-0.669	-1.692*
Quarter 2	-0.237	-0.514	-0.240	-0.584
Quarter 3	-0.605	-0.822	-0.701	-1.213
Quarter 4	-0.069	-0.106	-0.136	-0.213
Quarter 5	0.231	0.413	0.630	0.824
Quarter 6	0.473	0.511	0.264	0.317
Quarter 7	0.168	0.236	-0.148	-0.243
Quarter 8	1.107	1.852*	0.986	1.409

Notes: * indicates $p < 0.1$