Journal of Health Economics 42 (2015) 115–124

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

Does women's education affect breast cancer risk and survival? Evidence from a population based social experiment in education

Mårten Palme[a,*], Emilia Simeonova[b]

[a] Department of Economics, Stockholm University, SE-106 91 Stockholm, Sweden
[b] Johns Hopkins University and NBER, 100 International Drive, Baltimore, MD 21202, United States

ARTICLE INFO

Article history:
Received 12 July 2013
Received in revised form 30 October 2014
Accepted 5 November 2014
Available online 22 November 2014

JEL classification:
I12
I18

Keywords:
Education gradient in health
Schooling reform
Breast cancer

ABSTRACT

Breast cancer is a notable exception to the well documented positive education gradient in health. A number of studies have found that highly educated women are more likely to be diagnosed with the disease. Breast cancer is therefore often labeled as a "welfare disease". However, it has not been established whether the strong positive correlation holds up when education is exogenously determined. We estimate the causal effect of education on the probability of being diagnosed with breast cancer by exploiting an education reform that extended compulsory schooling and was implemented as a social experiment. We find that the incidence of breast cancer increased for those exposed to the reform.

1. Introduction

Worldwide, breast cancer is the most common cancer and the leading cause of cancer death among women. In the year 2008 alone, 2.6 women were diagnosed with breast cancer every minute across the globe. More than 52 women died of breast cancer every hour in the same year (Ferlay et al., 2010). These aggregate numbers mask large differences in trends in breast cancer incidence and mortality across the developed and developing economies. Historically, western societies have experienced a heavier burden of the disease, however in the last couple of decades the incidence and mortality from breast cancer has been on the rise in developing countries (Althius et al., 2005). While it is plausible that this rise is due to increased screening and better medical and vital records keeping, some have argued that more affluent societies and westernization also contribute to these recent trends (ibid).

Breast cancer in women is one of the rare health conditions that exhibit a positive incidence gradient with socio-economic status (SES), and in particular with attained education (see e.g. Hemminki and Li, 2003, 2004; Lund and Jacobsen, 1991; Hussain et al., 2008). This is in stark contrast with the well-documented negative association between education and all-cause mortality and with the positive effects of education on health-promoting behaviors (Cutler and Lleras-Muney, 2006, 2011 for reviews of the literature). Frequently diagnosed cancers of the female reproductive organs, such as cervical cancer, show the opposite, negative association with education in correlational studies (Baquet et al., 1990). Part of the observed positive correlation between education and breast cancer could be due to more frequent screening and more adequate response to risk factors among the better educated (Lange, 2011). Still, environmental and social factors could also affect breast cancer risk and survival. A recent report by the Interagency Breast Cancer and Environmental Research Coordinating Committee (IBCERCC) in the US forcefully argues that research on the causes of increased breast cancer risk and consequently on increased prevention is of first order importance in designing public health strategies to contain the disease.[1]

A key question on the etiological background to the link between education and the incidence of breast cancer is whether the relation is made up by life style factors, such as delayed childbearing,

* Corresponding author. Tel.: +46 8 16 33 07; fax: +46 8 15 94 82.
E-mail addresses: Marten.Palme@ne.su.se (M. Palme), Emilia.Simeonova@gmail.com (E. Simeonova).

http://dx.doi.org/10.1016/j.jhealeco.2014.11.001
0167-6296/

[1] "Prioritizing Prevention" Summary of Recommendations of the Interagency Breast Cancer and Environmental Research Coordinating Committee (http://www.niehs.nih.gov/about/assets/docs/ibcercc_full.pdf).

that may be acquired along with prolonged education, or if it can be attributed to factors and individual characteristics correlated with both educational attainments and the probability to get breast cancer. The most common research strategy used in epidemiological studies is to add confounders that are known to be associated with educational attainments and potentially etiologically related to breast cancer, such as delayed childbearing in a regression framework, and to investigate if the correlation remains (see e.g. Braaten et al., 2004; Danø et al., 2004; Heck and Pamuk, 1997). There are at least two problems with this strategy. First, there is an identification problem. Most confounders, such as fertility behavior, are likely to be endogenous to educational attainment. This means that it is still not clear if including them in the regression makes up for a causal relation with education, or if they just proxy individual characteristics correlated with educational attainments. Second, adding independent variables in a regression would in most cases aggravate the downward bias from measurement errors (see e.g. Greene, 2003).

An alternative strategy to analyze this research question is to use exogenous variations in educational attainments created by natural experiments. A number of influential studies have used this research strategy to study the relationship between education and measures of general health. Lleras-Muney (2005), Oreopoulos (2006) and Clark and Royer (2012) use variation induced by changes in compulsory schooling legislations in the US and the UK as a source of exogenous variation in education. Spasojevic (2010), Meghir et al. (2012) as well as Lager and Torssander (2012) investigate the health consequences of the introduction of comprehensive school reform in Sweden. An interesting related question is whether the health effects of education vary by gender[2] and diagnosis.

In this paper we investigate whether there is a causal effect of education on the incidence and mortality from breast cancer in the population of women born in Sweden between 1940 and 1957 who survived until at least 1985. We make use of a compulsory schooling reform that increased the number of compulsory years of education from 7 or 8 depending on municipality to 9 years nationwide. We also compile a unique nationally representative dataset from various Swedish national data registries, including the Swedish Cancer Registry.

The Swedish setting is particularly well suited to study how education affects the incidence of a "welfare disease" such as breast cancer in women for several reasons. First, Sweden is ethnically and racially homogenous, especially in the cohorts under study. This reduces potential omitted confounders that could correlate both with the hereditary genetic make-up and the SES of some ethnic or racial subgroups. Second, health care is free at the point of access and the Swedish government provides free universal health insurance. Disparities arising from differential access to care due to financial constraints are unlikely to play a role in the Swedish setting. Breast cancer screening covers the entire female population in the critical ages and is free of charge. The screening program was adopted nation-wide in 1986 after the first results from the Swedish mammography trials became available (Tabar et al., 1985). The take up rate of this screening program after the first invitation to screen is about 80 percent (see e.g. Hussain et al., 2008). Third, the Swedish Cancer Registry is the oldest cancer registry and one of the best in terms of data quality and accuracy in the world today.

The closest study we are aware of is by Glied and Lleras-Muney (2008) who use the Surveillance, Epidemiology and End Results Program (SEER) data to estimate the effects of technological progress on cancer deaths by education, relying on US compulsory schooling laws for exogenous variation in educational attainment. They find that conditional on technological progress, extra education reduces overall cancer mortality in men, but not in women. Excluding cancers of the reproductive system, inclusive of breast cancer, makes the estimated effects for men and women consistent. The authors do not specifically test for the effects of education on survival from reproductive system cancers in women, relying on the findings in the medical literature we discuss above.

This study finds that attaining higher levels of education increases the risk of being diagnosed with breast cancer in women, confirming the results obtained from purely correlational studies. However, we also find that this heightened probability of diagnosis is later followed by an elevated probability of death from breast cancer among better educated women. Further, we investigate the potential role of fertility decisions, which has been pointed out as the mechanism linking education and the incidence of breast cancer. We find no convincing evidence in favor of this hypothesis. The curious association between education and the most common cancer diagnosis in women appears to be affected by qualities, behaviors, and risk factors acquired in the process of obtaining more education, rather than pre-existing characteristics that predispose some women to both get more education and be diagnosed with the disease.

2. The comprehensive school reform

2.1. The Swedish school system before and after the reform

Sweden implemented a compulsory schooling reform as a social experiment between 1949 and 1962. Prior to the implementation of the reform, pupils attended a common basic compulsory school (*folkskolan*) until grade six. After the sixth grade pupils were selected to continue either for one or, in mainly urban areas, two years in the basic compulsory school, or to attend the three year junior secondary school (*realskolan*). The selection of pupils into the two different school tracks was based on their past academic performance, measured by grades. The pre-reform compulsory school was in most cases administered at the municipality level. The junior secondary school was a prerequisite for the subsequent upper secondary school, which was itself required for higher education.

In 1948 a parliamentary committee proposed a school reform that implemented a new nine-year compulsory comprehensive school.[3] The reform had three main elements:

1. An extension of the number of years of compulsory schooling to 9 years in the entire country.
2. Abolition of early selection and tracking based on academic performance. Although pupils in the comprehensive schools were able to choose between three tracks after the sixth grade – one track including vocational training, a general track, and an academic level preparing for later upper secondary school – they were kept in common schools and classes until the ninth grade.
3. Introduction of a national curriculum. The new curriculum replaced the pre-existing curriculum which varied between municipalities.

[2] Clark and Royer (2012) as well as Meghir et al. (2012) investigate for differential effects of education by gender and find inconclusive evidence. Gathman et al. (2012) analyze a number of compulsory schooling reforms in Europe and find diverging effects of education on mortality by gender.

[3] We offer a brief description of the main parts of the Swedish comprehensive school reform. The school reform and its development are described in Meghir and Palme (2003, 2005), and Holmlund (2007). For more detailed reference on the reform, see Marklund (1981).

2.2. The social experiment

The social experiment with the new comprehensive nine-year compulsory school started during an assessment period between 1949 and 1962, when the final curriculum was decided.[4] The proposed new comprehensive school system, as described above, was introduced in municipalities or parts of city communities, which in 1952 numbered 1055. The cohorts included in our empirical analysis, born between 1940 and 1957, cover the entire period of implementation of the comprehensive school. In 1962 it was decided that the new comprehensive school would become the standard education in Sweden. The last class that graduated from the old schooling system did so in 1970.

The selection of municipalities into the new comprehensive school was not based on random assignment. Still, the decision to select the areas was based on an attempt to choose locations that were representative for the entire country, both in terms of demographics as well as geographically. At first the National Board of Education contacted the municipalities, or sometimes they themselves applied to participate. From this pool of applicants a "representative" sample of municipalities was chosen. Municipalities could elect to implement the comprehensive school starting with first or fifth grade cohorts. Once the grade of implementation was fixed, all individuals from the cohort immediately affected and all subsequent cohorts went to comprehensive school. The older cohorts continued in the per-reform school.

Meghir and Palme (2005) and Holmlund (2007) study the effect of the comprehensive school reform on educational attainments.[5] The Meghir and Palme (2005) estimates for their entire sample are 0.252 additional years for males and 0.339 years for females; for low SES persons the estimates are 0.3 extra years for males and 0.512 for females.[6] Holmlund has estimates in the range 0.21–0.61 additional years of schooling for men and 0.13–0.44 for women.

3. Data

This is a population-level study. We match data from the Swedish Cancer Register to Swedish population register data, the 1990 Swedish Education register, and the Cause of Death register. The population register contains information on the parish of birth for all individuals born in Sweden in the 1940–1957 cohorts. We use this register to assign municipality of birth for all women in the cohorts affected by the schooling reform. The municipality of birth is then used to assign the year in which the reform was implemented in that locality, and the reform treatment status to different cohorts of women who were born in the municipality. Note this means that all estimated effects are of the "intention to treat" type, but we avoid potential bias coming from selective migration. Holmlund (2007) offers a detailed exposition of the exact matching algorithm used.[7]

The Cause of Death register contains information on the date of death and the principal cause of death. The Census data provide information on the date of birth and the number of children born to the women in the 1940–1957 cohorts. We use this information to assign age at first childbearing and the total completed fertility per woman. The multi-generational register is used to link women in the sample to their parents. We then use the Education Register for the parents to determine the level of education of each woman's father. Fathers who had more than the basic required (7 years) education are considered highly educated.

All women who died of breast cancer as a primary cause of death were found in the Cancer Register as having been previously diagnosed with the disease. We record all diagnoses and deaths until 2006. The Swedish Cancer Register is the oldest Cancer Register in the world and contains detailed information on all incidences of cancer diagnosis in Sweden. It is compiled from compulsory cancer diagnosis registrations by physicians, cytologists and pathologists and covers close to 100% of all cancer diagnoses in Sweden (Swedish National Board of Health and Welfare, 2006). Studies of the accuracy of the Cancer Register have shown that cases of breast cancer are the most reliably reported cancer diagnosis in the Register, with under-reporting rates of less than 1.1% of all cases diagnosed within the reporting year (Barlow et al., 2009). Importantly, the exact date of every diagnosis is recorded, and the data can be linked to the population registers through a unique person ID.

In the empirical analysis we use the population of all women born in Sweden between 1940 and 1957 and surviving until at least 1985. We exclude 414,214 women with missing parental education background and use the remaining sample of 562,814 women. Of those, 19,736 women were diagnosed with breast cancer after 1984. Of those who were diagnosed, 2370 women died, and breast cancer was noted as the cause of death on their death certificate. Another 401 of the women diagnosed with breast cancer after 1984 died from a different main diagnosis.[8]

Table 1 summarizes the main explanatory and control variables used in the analysis. The mortality data start in 1985 and include the exact date of death and the main cause of death as recorded in the death certificate. We restrict the time of first diagnosis to be after 1984 in order to avoid selection of women who were diagnosed previously and survived until the period after 1984. The women in our sample were aged between 28 and 45 in 1985 and (those surviving) between 49 and 66 in 2006. As a percent of total female mortality, breast cancer mortality peaks between ages 40 and 60 at around 15% of total deaths in the age group. This implies that we are capturing the interval in women's lives during which they are most likely to be affected by breast cancer (as opposed to another lethal disease). Aggregate mortality in Sweden is very low at ages below 45 at 6 per 1000 (from data), and breast cancer mortality is even lower at 1 per 1000 (from data). A back of the envelope calculation suggests that we are potentially missing at most 100 deaths from breast cancer that may have occurred in our study population before 1985.[9] This is a very small part of the total number of breast cancer deaths in the sample – less than 5%.

Several differences in the raw means between women of high and low SES family backgrounds are worth discussing. Unsurprisingly, on average women of higher SES obtained more years of education. They are less likely to have had any children and the average age at first childbearing in this group is about two years higher. High SES women are also more likely to have been

[4] The official evaluation was mainly of administrative nature. Details on this evaluation are also described in Marklund (1981).

[5] Holmlund (2007) does not have individual treatment status and imputes it from municipality of residence in 1960.

[6] Note that Meghir and Palme (2005) use the exact reform assignment from the school registries for a random subset of the cohorts born in 1948 and 1953. Their estimates are free of measurement error in the reform assignment variable.

[7] We are grateful to Helena Holmlund for sharing her algorithm with us.

[8] The distribution of the main causes of death among those women is: 61 women died from ovarian cancer; 45 from lung cancer; 22 from AMIs; 15 from pancreatic cancer; 12 died from colon cancer; 9 from melanoma; 15 from unknown causes. The remaining 222 deaths are distributed across more than one hundred different causes.

[9] Assuming mortality at missing ages is equivalent to mortality at those ages among observed cohorts; assuming also that cohort sizes at different ages are similar over time, which gives an upper bound estimate since demographic trends led to steady cohort size increase between 1940 and 1957.

Table 1
Main explanatory and outcome variables in interest. Standard deviations are reported in square brackets under the mean. *P*-values of tests of differences in means between high and low SES background women are also reported.

Father's education	Low FE		High FE		*P*-value diff
	Obs	Mean	Obs	Mean	
Variable					
Years of education	360,240	11.155 [2.805]	180,612	12.823 [3.063]	0
No children (nulliparous)	372,894	0.11 [0.315]	187,702	0.134 [0.34]	0
Age at first childbearing	331,164	26.5 [5.4]	162,587	28.1 [5.5]	0
Age at diagnosis	12,723	50.196 [6.581]	6654	49.513 [6.511]	0.088
Death year – year of diagnosis	1538	4.83 [4.55]	724	5.3 [4.63]	0
Diagnosed with breast cancer	372,894	0.035 [0.183]	187,702	0.036 [0.186]	0.012
Died from breast cancer	372,894	0.004 [0.065]	187,702	0.004 [0.063]	0.16

diagnosed with breast cancer, to have received the diagnosis at an earlier age and, conditional on dying from breast cancer, to have lived longer between their initial diagnosis and the time of death. There is no significant difference in the probability of death from breast cancer by SES background. These facts suggest that either (1) higher SES women are more likely to have been diagnosed earlier or that (2) higher SES women received better treatment, or both. The differences in age at diagnosis appear in favor of the first hypothesis, but we cannot draw any firm conclusions based on this evidence.

Here it is important to consider the importance of breast cancer screening for early diagnosis and treatment. The large clinical trials that produced evidence on the beneficial effects of mammography, were done in Sweden in the 1970s and 1980s (see Tabar et al., 1985). Thus, policy makers in Sweden were quite aware of the importance of breast cancer screening at the time our study period begins. After the first results of the randomized trials came out, the National Board of Health and Welfare issued guidelines in 1986 recommending that the county councils invite women ages 40–54 years to screening every 18 months and women ages 55–74 years every second year. Thus, national service screening with mammography was initiated in 1986. Local health administrations are in charge of running the screening programs. All women of eligible ages receive a letter giving a specific date and time for a mammography examination. Failure to attend the scheduled examination or re-schedule the appointment results in a second invitation up to six consecutive invitations. A regional case study from Uppsala reports that of the 46,041 eligible women only 5.6% never attended after six attempted appointments. Non-attenders tend to be older (over the age of 60), foreign born and single. Note that all foreign-born women residing in Sweden are excluded from our sample by construction. Interestingly, the relationship between education and the probability of non-attendance is u-shaped, with women finishing high school, some college, and college more likely to attend than those with professional education or high school drop outs (Lagerlund et al., 2002).

Breast cancer is a common killer in our sample. In this relatively young population, 15% of all deaths are due to breast cancer. Cardio-vascular diseases account for an extra 13.5% of total mortality. Deaths from other cancers are responsible for another 33%. In total, cardio-vascular and cancer-related mortality account for close to two-thirds of all female deaths in the sample cohorts.

4. The relation between educational attainment and breast cancer incidence and mortality

We start the analysis by documenting correlations between the years of attained education and socio-economic background and the probability of diagnosis and death from breast cancer in Sweden. Table 2 presents the estimates. We use all available observations to maximize power. Coefficients and standard errors are multiplied by 1000 for better presentation. Women with an extra year of education are 3 percent (evaluated at the mean incidence of breast cancer diagnosis in the population) more likely to have been diagnosed with breast cancer than their less educated peers. The correlation with high socio-economic family background is larger – even if we control for years of attained education, women who were born in better-off families are 7 percent more likely to receive a breast cancer diagnosis.

The effect of education on deaths from breast cancer is not as clear. None of the education coefficients attain statistical significance at the 10% level even though the coefficient on years of education implies a negative correlation both with and without

Table 2
Correlations between years of educational attainment and diagnosis/death from breast cancer in women.

Panel A	(1)	(2)	(3)
Diagnosis			
Years of schooling	0.96*		0.89*
coef * 1000	(0.09)		(0.08)
High SES		2.99*	1.92*
coef * 1000		(0.54)	(0.56)
Observations	541,135	560,596	541,135
Mean incidence per 1000	34.1	34.1	34.1
R-squared	0.006	0.007	0.010
Empirical model	Linear Prob	Linear Prob	Linear Prob
Panel B	(1)	(2)	(3)
Death from breast cancer			
Years of schooling	−0.01		−0.02
coef * 1000	(0.01)		(0.01)
High SES		−0.09	0.12
		(0.18)	(0.09)
Mean deaths per 1000	4.1	4.1	4.1
Observations	541,135	560,596	541,135
Empirical model	Linear Prob	Linear Prob	Linear Prob

Note: Robust standard errors in parentheses; *Significant at 1%; SE clustered on the municipality of birth level.

controls for parental SES background. Coupled with the evidence on the higher incidence of diagnoses among the more educated women, this suggests that conditional on being diagnosed with breast cancer, more educated women are more likely to survive. This is consistent both with evidence that educated people are more adept at using new medical technologies (Glied and Lleras-Muney, 2008; Lichtenberg and Lleras-Muney, 2005) and with earlier diagnosis and earlier treatment in higher SES background women. The table of means shows supportive evidence for the latter explanation. Even though the mortality point estimate suggests that education has a positive effect on survival, the precision of the estimates is not high enough to make any strong conclusions.

The corresponding Cox proportional hazard estimates are: in the full sample one year of schooling reduces the probability of death from breast cancer by a statistically insignificant 1.6% relative to the mean (SE 0.0197), which is a larger estimate than the LP estimate evaluated at the mean (0.03%); the Cox estimate of high SES in model 2 is a statistically insignificant decrease of 2.4% relative to the mean, not too far from the LP estimate of 2.2% relative to the mean.

5. Empirical specification

We use two main types of outcomes in the empirical analysis. When we consider breast cancer mortality, we use the binary mortality outcome and the time to death as the outcome variables. When we study the incidence of breast cancer, we use a binary outcome variable equal to one if the woman was ever diagnosed with breast cancer after 1984 and zero otherwise. We use the same identification strategy for the effect of the reform for both types of outcomes. If the reform would have been randomly distributed among Sweden's 1000 or so municipalities we could have simply compared the outcomes in the treated and non-treated municipalities conditional on year of birth. However, as has been discussed in previous studies (see e.g. Meghir et al., 2012), this was not the case. Therefore, we will control for both birth cohort and municipality of birth. We start with the following latent variable specification:

$$y^*_{i,m,t} = \alpha + \beta_1 R_{i,m,t} + \gamma_1' T_i + \gamma_2' M_i + \varepsilon_{i.m.t}, \quad (1)$$

where i, m and t are sub-indices for individual, municipality and birth cohort, respectively; y^* is a latent variable for health status; T is a vector of dummy variables for year of birth; M is a corresponding vector of dummy variables for municipality of birth; finally, ε is an individual random disturbance.

The key identifying assumption is that the distribution $f(\cdot)$ of ε does not depend on the assignment to reform treatment, conditional on cohort and municipality. In practice we impose the stronger assumption that the distribution of ε is independent of all right hand side variables. It is important to note that the reform assignment in this analysis depends on the municipality of birth, rather than the municipality of schooling. On the one hand, this means that the estimates are of the "intention-to-treat" type. On the other hand we avoid selection issues coming from differential (and potentially endogenous) mobility.[10]

For the binomial outcome breast cancer diagnosis, we use linear probability models. The reason for using a linear probability model, rather than e.g. logit and probit, which restrict the probabilities in the [0, 1] interval and relax the linearity assumption, is computational convenience, since all models include about 1000 municipality indicator variables in addition to the 17 birth cohort dummies. For relatively small treatment effects, when both approaches have been used in a similar context, the results are almost identical.[11]

We also use linear probability models as one of two methods of estimating the probability of death from breast cancer. The linear probability model is handy because it can efficiently estimate a large number of dummy coefficients in specifications where we also include municipality-specific time trends. We complement the linear probability estimates with estimates from Cox semi-parametric proportional hazard models. For the time to death from breast cancer outcomes we use Cox proportional hazard models of this type:

$$I_{1,i,m,t}(r|R_{i,m,t}, T_i, M_i) = I_0(r)\exp\{\alpha + \beta_1 R_{i,m,t} + \gamma_1' T_i + \gamma_2' M_i\}, \quad (2)$$

where r is exposure time and $I_0(r)$ is the baseline hazard. This model is semi-parametric in the sense that no functional form assumption is imposed on the base line hazard. Importantly, when we consider the hazard of death from breast cancer, we consider only deaths from breast cancer as terminal event. Thus, all women who died from causes other than breast cancer are considered still living at the end of the observation window. Prior research has found that the compulsory schooling reform did not significantly affect life expectancy for (high and low SES) Swedish women (Meghir et al., 2012). Moreover, the age at first diagnosis in this sample is fairly young. These two facts suggest that a competing risks phenomenon is an unlikely explanation for our estimates. Nevertheless, as Honoré and Lleras-Muney (2006) show that decreasing cardio-vascular disease mortality in the US contributed to a steady (non-declining) cancer mortality rate between the 1970s and 2000s we construct Peterson bounds on our estimates taking into account the association between cardio-vascular and cancer mortality risks. The assumption here would be that obtaining extra education, while reducing the likelihood of cardio-vascular mortality, indirectly increases the likelihood of breast cancer mortality. To control for unobserved differential trends that might affect municipalities differently depending on the timing of the education reform, we include linear trends by year of reform implementation. All municipalities that implemented the reform in the same year are assigned the same linear trend. The empirical results section reports the results from these preferred specifications.

Table 3 demonstrates the effects of being exposed to the schooling reform on the number of years of attained education for Swedish women of affected cohorts. We first show the effects on the entire sample and then split the sample according to the education level of the father. We expect that the education reform affected children from low SES families more as they were more likely to drop out of school earlier. The results confirm that women from low SES backgrounds increased their education by more than those from high SES backgrounds. The reform resulted in an average increase of 1.8 months of schooling for girls coming from relatively disadvantaged backgrounds. The corresponding estimate for the high SES group is about one third of the size and does not attain statistical significance at the 10% level. We present estimates from models including linear time trends grouped by year of implementation and from specifications including municipality-specific linear time trends. The estimated coefficients are very similar, which is reassuring that unobserved municipality-specific changes coincidental with reform implementation are unlikely to bias our results.

[10] Meghir and Palme (2005) show, however, that 90.1 percent have the same reform assignment based on predictions from their municipality of birth as their municipality of schooling; 5.3 percent moved from reform to non-reform municipalities; 4.6 moved in the other direction.

[11] See for example Meghir et al. (2011).

Table 3
The effect of education reform on women's educational attainment in years of education.

Father's education	All		Low FE		High FE	
	(1)	(2)	(3)	(4)	(5)	(6)
Reform	0.119*	0.106*	0.149*	0.14*	0.055	0.043
	(0.029)	(0.023)	(0.022)	(0.021)	(0.035)	(0.033)
Mean years of education	11.7	11.7	11.1	11.1	12.8	12.8
Linear trends by year of reform implementation	Yes		Yes		Yes	
Municipality trends		Yes		Yes		Yes
Observations	540,852	540,852	360,240	360,240	180,612	180,612
R-squared	0.041	0.045	0.041	0.045	0.024	0.031

Note: Robust standard errors in parentheses; SE clustered on the municipality of birth level. +Significant at 10%; **significant at 5%; *significant at 1%.

Table 4
Educational reform and the risk of diagnosis and death from breast cancer.

Father's education	All		Low		High	
Diagnosis						
Reform coef * 1000	1.5**	1.71**	1.6	1.92+	1	1.1
	(0.78)	(0.775)	(1.1)	(1.025)	(1.5)	(1.47)
Mean dep var * 1000	34.1		33.7		35	
Linear trends by year of reform implementation	Yes		Yes		Yes	
Municipality trend		Yes		Yes		Yes
Observations	560,596	560,596	372,894	372,894	187,702	187,702
R-squared	0.006	0.0075	0.006	0.009	0.010	0.015
Death from breast cancer						
Reform coef * 1000	0.63**	0.66**	0.68+	0.74+	0.65	0.6
	(0.3)	(0.33)	(0.4)	(0.44)	(0.9)	(0.87)
Mean dep var * 1000	4.1	4.1	4.2	4.2	3.9	3.9
Linear trends by year of reform implementation	Yes		Yes		Yes	
Municipality trend		Yes		Yes		Yes
Observations	560,596	560,596	372,894	372,894	187,702	187,702
R-squared	0.002	0.004	0.003	0.0061	0.006	0.010

Note: Municipality fixed effects included in all specifications; birth cohort dummies included in all specifications; robust standard errors in parentheses; standard errors clustered on the municipality of birth level; +significant at 10%; **significant at 5%; *significant at 1%.

6. Results

6.1. Main findings

We next turn to the effects of the compulsory education reform on breast cancer incidence and death. Since we know from previous research (see e.g. Meghir and Palme, 2005) that the reform had very different effects on later-life economic wellbeing depending on parental SES, we run separate regressions by women's family SES background. It is important to note that since the reform was not limited to simply increasing the number of compulsory years of education, but had additional elements, the results that follow are not directly comparable with the education correlations presented in Table 2. The results on how reform treatment affected the probability of diagnosed breast cancer are shown in the top panel of Table 4 and on mortality from the disease in the bottom panel. We present estimates with year of implementation specific linear trends for easy comparison with the semiparametric Cox estimates, as well as results from linear probability models including municipality-specific linear trends. We multiply all coefficients and standard errors by 1000 for ease of presentation.

There is a significantly positive effect of reform assignment on the probability of being diagnosed with breast cancer in the full sample. Although the precision in the estimate in the low SES subsample is somewhat inferior, it is obvious that the effect is attributable to the group originating from low SES families, who experienced the largest effect of the education reform. The point estimate of the magnitude of the effect suggests a 1.5 percentage point elevated risk, which is somewhat more than the correlation estimate corresponding to one year extra year of education although the precision is not sufficient to make any definite conclusion.[12]

The linear probability estimation results in Table 4 also show that the reform causes a significant increased risk of mortality with breast cancer as a primary cause of death.[13] That is, the expected improvement in the effect of better response to cancer treatment from more education was not sufficient to offset the increased risk of being diagnosed with breast cancer.

In addition to the linear probability models we obtained Cox PH mortality estimates stratified at the municipality level. The Cox semi-parametric model imposes fewer restrictions on the estimates, however it suffers from severe incidental parameters problems with a large number of dummy variables, such as would be included in a specification including municipality specific linear trends. That is why we ran the Cox estimations with linear trends by year of implementation only, so these estimates are comparable to the linear probability coefficients reported in columns (1), (3), and (5). The hazard is stratified by municipality of birth, allowing for potentially different underlying breast cancer mortality hazards by municipality. The Cox estimates are as follows: full

[12] If we consider only the reform's effect on schooling attainment, we would multiply the reform estimates from Table 4 by 1/(estimated change in years of education from Table 2), resulting in much larger estimates of the effect of an additional year of education than what is obtained in the correlations reported in Table 3. We emphasize, however, that using the reform as an IV for years of attained education is most likely flawed, as the reform contained additional elements that challenge the exclusion restrictions.

[13] In the analyses we exclude all women who have received a diagnosis of breast cancer pre-1985 to avoid selection bias. This is because our mortality data start in 1985 and survival following a breast cancer diagnosis could be related to the reform.

sample coefficient 0.18+ (SE 0.1), which is very similar to the LP estimate evaluated at the mean (15% increase); low father's education sample estimate 0.2 (SE 0.15) – again very similar to the LP coefficient estimated at the mean – a 16% increase.

6.2. Competing risks

A potentially important concern in analyzing mortality by different causes has been raised by Honoré and Lleras-Muney (2006). Technological progress in medicine or any other factor that affects the treatment or detection of certain diseases would affect the probability of death from related diagnoses but also the probability of death from other conditions, which pose "competing risks". In essence, failure to die from one condition at a given age increases the probability of death from another condition. Honoré and Lleras-Muney (2006) show in particular that cardio-vascular (CVD) and cancer deaths in the US are related in this manner. Improvements in the treatment of CVD led to decreased mortality from CVD but also to increased mortality from cancer compared to the counterfactual. This is important in our setting because education may have affected the early detection and proper treatment of CVD, leading to a reduction in the probability of death from CVD. Through the competing risks channel, this reduction may have increased the probability of death from breast cancer. To examine this hypothesis, we first estimate the effects of being exposed to the reform on CVD mortality and compute bounds for our estimates.

The probability of death from CVD is reduced by reform treatment by 0.53% (Cox estimate 0.99472, CI 0.8353–1.1845) in the full sample. In the subsample of low SES background women, the reform treatment leads to a 2.7% decrease in CVD mortality (Cox estimate 0.97304, CI 0.80374–1.178). Assuming that everyone who did not die from CVD died from breast cancer, we compute a lower bound on our breast cancer mortality estimates. The education reform increases the probability of death from breast cancer or CVD by 9.8% (hazard ratio 1.098, CI from 0.997 to 1.209). Thus even if the entire reduction in CVD mortality is translated into breast cancer mortality, we still find a positive effect of the reform on the (combined) mortality rate, even though it is about half the size of the effect we obtain when we assume the risks are unrelated (18%).[14]

6.3. Parallel trends assumption

Our difference-in-differences analysis relies on the assumption of parallel trends in the incidence of diagnosed breast cancers before and after the cohort affected by the reform in each municipality. We implement two different tests of this assumption. First, Fig. 1 plots the conditional marginal effects of exposure in the 6 cohorts pre-implementation to 6 cohorts post-implementation sample. The reference cohort is the one born 2 years before the first treated cohort. The regressions control for municipality and year of birth fixed effects, as well as municipality group by year of implementation linear trends. As the figures demonstrate the conditional probability of diagnosis and death is not significantly different from zero in cohorts born pre-implementation. There is however a sharp increase in the probability that starts with the cohort right before the first fully treated cohort and levels off at a new and increased level with the second fully treated cohort (1 year after year zero of the implementation in the figures below).

Second, we performed a number of placebo tests in which we pretend that the reform was implemented earlier or later than the actual implementation year. The placebo treatment groups are defined by falsely assigning treatment to women born 6, 4 and 2 years before the first fully treated cohort, as well as 2, 4 and 6 years after the first cohort. In the first arrangement women who were not treated receive false treatment status. In the latter arrangement we pretend that women who were (actually) treated and were born 2, 4, and 6 years from the first treated cohort were not treated. Thus, in this set-up treated women receive false untreated status. We present all these tests together in Fig. 2.

Every estimate is obtained from a separate regression including cohort and municipality fixed effects, as well as year of implementation linear trends. The regressions assigning treatment to untreated cohorts include only women from untreated cohorts. Similarly, the regressions assigning non-treatment to treated cohorts include only treated women. As the figures demonstrate, the largest in absolute value and only statistically significant effects are estimated when we assign the correct treatment values. Further, there are no particular discernible patterns, suggesting that there is nothing that systematically biases our estimates.

6.4. Changes in fertility behavior as a possible mechanism behind the results

The causal estimates confirm the positive correlations between education and the probability of breast cancer diagnosis. Medical studies have pointed to several channels that might contribute to these findings (see Nechuta et al., 2010 for a recent review). For two of these – the inverse relation between educational attainments and completed fertility as well as the positive relation between education and age at first birth – we have information in our data set allowing us to analyze how these two outcomes were affected by the schooling reform.

As a background to this analysis, Table 5 shows associations between attained education and women's fertility behavior in Sweden using the same population we analyzed in the mortality regressions. Column (1) reports the correlation between year of schooling and the probability of having no children; column (2) displays the relation between year of schooling and number of children; finally, column (3) The shows the correlation between years of schooling and age at first child.

As can be seen in Table 5, there is a statistically significant relation between years of schooling and each of the three outcomes under study. The point estimates suggest that one additional year of schooling is associated with a 0.003 increase in the probability of having no children; 0.013, or an about 0.8 percent, fewer children; and, finally, almost half a year older age at first birth.

In Table 6 we turn to analyzing the effect of schooling reform on the same set of outcomes as those analyzed in Table 5. None of the point estimates attain statistical significance. Comparing the estimates for the effect of the reform with the correlations shown in Table 5, it is evident that the precision in the reform effect estimates for the probability of having no children as well as the total number of children is too low to enable us to reject

[14] A separate issue emerges if we consider testing for the effect of the reform on deaths from breast cancer as one of a series of multiple mortality tests we could perform, including the reform effect on death from CVD and death from other causes. We performed an adjustment procedure to calculate the q-value, which is the P-value of the test adjusted for the false discovery rate. This methodology was developed by Storey and co-authors and software was created by Dabney and Storey (Storey, 2002; Storey and Tibshirani, 2003). We picked a π_0 of 1 and an FDR threshold of 0.05. The q-value on the reform coefficient in the first linear probability regression of breast cancer mortality is 0.096 (the P-value is 0.032); the q-value on the reform coefficient in a linear probability regression with binary outcome "death from any other condition" is 0.141 (P-value 0.95); the q-value on the reform coefficient in a linear probability regression with binary outcome "death from CVD" 0.79 (P-value 0.79). While a P-value threshold of 0.1 implies that 1 in every 10 tests will be a false positive, a q-value threshold of 0.1 implies that one in every 10 positive tests will be a false positive. Even after a conservative adjustment for multiple hypotheses testing, we obtain a reform effect that is still significant at the 10% level.

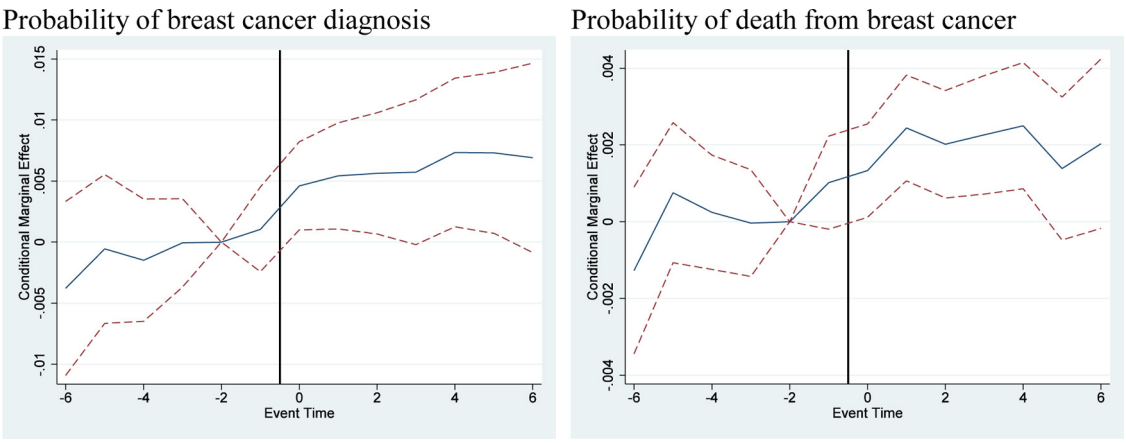

Fig. 1. Probability of breast cancer diagnosis and death from breast cancer among cohorts of women born close to the first cohort affected by the reform. *Note*: Conditional marginal effects plotted in solid line, 95% confidence intervals in dashed lines. The omitted category is women born 2 years before the first cohort that was affected by the reform. Cohort and municipality fixed effects included in the regressions, as well as municipality group by year of implementation-specific linear trends.

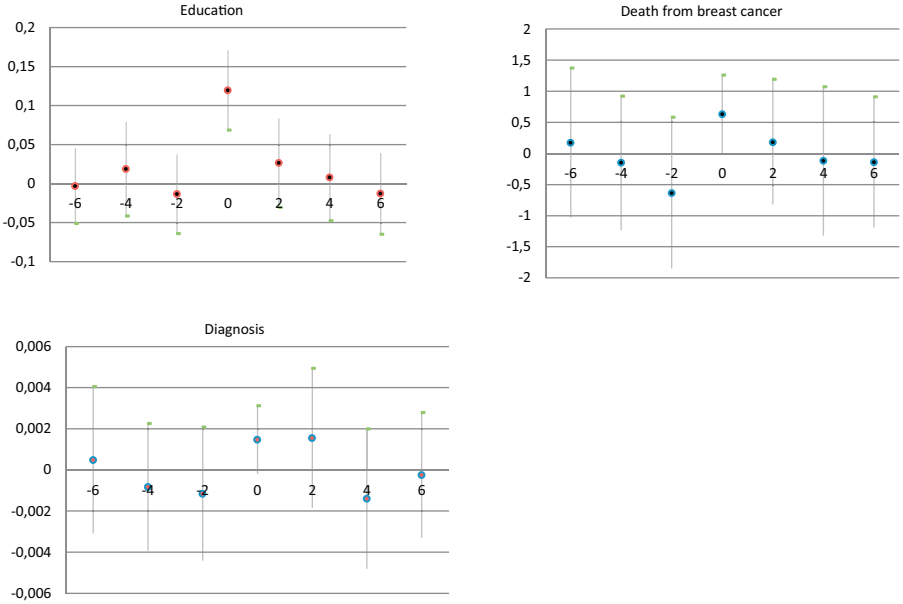

Fig. 2. Placebo tests assigning treatment status to untreated or untreated status to treated cohorts.

the hypothesis that the effects are the same as for one additional year of schooling. However, for the age at first child outcome, the point estimate is very different and the precision sufficient to allow us to reject that the effect is as large as the almost 0.5 years as suggested by the result in Table 5. The upper confidence limit for a 95 percent confidence interval for the reform effect is as small as 0.025 for age at first birth, and so we can conclude that it is unlikely that the mechanism behind our result of the reform effect on cancer diagnosis incidence is through delayed childbearing among those who had children. For the other two outcomes, the precision is too low for any definite conclusions.

Table 5
Education and women's fertility behavior.

Father's education Outcome	All No children (1)	All Total fertility (number of children) (2)	All Age at first childbearing (3)
Years of schooling	0.0034*	−0.0134*	0.4608*
	(0.0002)	(0.0010)	(0.0038)
Mean outcome variable	0.11	1.7	27
Empirical model	Linear Prob	OLS	OLS
Observations	541,135	541,135	478,946
R-squared	0.0057	0.0057	0.0868

Note: +Significant at 10%; **significant at 5%; *significant at 1%; robust standard errors in parentheses clustered at the municipality of birth; linear trends by year of reform implementation included in all specifications; birth cohort dummies included in all specifications.

Table 6
Education reform and women's fertility behavior.

Father's education	All (1)	Low education (2)	High education (3)
Probability of not bearing a child			
Reform	0.0018	0.0011	0.0033
	(0.0018)	(0.0021)	(0.0028)
Mean outcome variable	0.11	0.11	0.134
Empirical model	OLS	OLS	OLS
Observations	560,596	372,894	187,702
R-squared	0.0050	0.0057	0.0082
Total fertility (number of children)			
Reform	0.0028	0.0034	−0.0018
	(0.0061)	(0.0073)	(0.0107)
Observations	560,596	372,894	187,702
R-squared	0.0045	0.0055	0.0077
Age at first childbearing			
Reform	−0.0445	−0.0464	−0.0426
	(0.0350)	(0.0364)	(0.0572)
Mean outcome variable	27	26.5	28.1
Empirical model	OLS	OLS	OLS
Observations	493,751	331,164	162,587
R-squared	0.0269	0.0225	0.0319

Note: +Significant at 10%; **significant at 5%; *significant at 1%; robust standard errors in parentheses clustered at the municipality of birth; linear trends by year of reform implementation included in all specifications; birth cohort dummies included in all specifications.

7. Concluding remarks

Numerous studies have shown that higher educational attainment is conducive to better health in the affected cohorts and their offspring (see review by Cutler and Lleras-Muney, 2011). Breast cancer is an exception to this rule in the sense that the incidence of diagnosed cases increases with education and it has therefore been labeled a "welfare disease". We show that this relation holds also as a response to an exogenous policy change that induced an increase in compulsory schooling. This result suggests that the relation between women's educational attainments and breast cancer is likely due to some characteristic or risk factor that is acquired as additional education is obtained, rather than some innate quality that is correlated with educational attainments. Many social and health behaviors fit these categories, such as the use of hormonal therapies and oral contraceptives, which have been linked to increased probability of breast cancer.

The reform may have made women more willing to participate in screening programs. Meghir et al. (2013) shows that cognitive skills were improved as a result of the reform, which, in turn, may make people more adequately aware of different risk factors (see e.g. Cutler and Lleras-Muney, 2006). Since participation in screening programs is not included in our data, we are not able to estimate this effect separately. However, given that about 80 percent of Swedish females participate in the nationwide screening program it is not likely that improved participation in the program makes up the entire effect. Further, superior screening among higher SES women by itself would not explain the elevated risks of death that we document. The findings in this study imply that the quality and availability of breast cancer screening and preventive health care must keep pace with improving educational opportunities for women world-wide.

We find no convincing evidence that fertility behaviors, often cited as a potential mechanism behind the higher incidence of breast cancer in educated women, were significantly affected by the reform and we could therefore not conclude that it is driving the result of elevated risk of breast cancer caused by the education reform. For delayed childbearing the precision in our estimates were sufficient for excluding it as a major mechanism behind our results on breast cancer incidence and mortality.

Epidemiological studies of the incidence of breast cancer have discussed the possibility that breast feeding, breast feeding duration and the duration of oral contraceptive use may affect the probability of breast cancer. Since we have not been able to find large enough dataset including these outcomes, we have not been able to explore their potentials as a possible mechanism behind our results and have to leave this for further research.

References

Althius, M., et al., 2005. Global trends in breast cancer incidence and mortality 1973–1997. International Journal of Epidemiology 34 (2), 405–412.

Baquet, C., Horm, J., Gibbs, T., Greenwald, P., 1990. Socio-economic factors and cancer incidence among Blacks and Whites. Journal of the National Cancer Institute 83 (8), 551–557.

Barlow, L., et al., 2009. The completeness of the Swedish Cancer Register – a sample survey for year 1998. Acta Oncologica 48 (1), 27–33.

Braaten, T., Weiderpass, E., Kumle, M., Adami, H.-O., Lund, E., 2004. Education and risk of breast cancer in the Norwegian-Swedish women's lifestyle and health cohort study. International Journal of Cancer 110 (4), 579–583.

Center for Epidemiology, 2006. Cancer Incidence in Sweden 2004. Center for Epidemiology, National Board of Health and Welfare, Stockholm.

Clark, D., Royer, H., 2012. The effect of education on adult mortality and health: evidence from Britain. American Economic Review 103 (6), 2087–2120.

Cutler, D., Lleras-Muney, A., 2006. Education and health: evaluating theories and evidence. In: House, J.S., Schoeni, R.F., Kaplan, G.A., Pollack, H. (Eds.), The Health Effects of Social and Economic Policy. Russell Sage Foundation, New York.

Cutler, D., Lleras-Muney, A., 2011. Education and Health (Unpublished manuscript).

Danø, H., Hansen, D.K., Jensen, P., Pedersen, J.H., Jacobsen, R., Ewertz, M., Lynge, E., 2004. Fertility pattern does not explain social gradient in breast cancer in Denmark. International Journal of Cancer 111 (3), 451–456.

Ferlay, J., Shin, H.R., Bray, F., Forman, D., Mathers, C., Parkin, D.M., 2010. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International Journal of Cancer 127 (12), 2893–2917.

Gathman, C., Jürges, H., Reinhold, S., 2012. Compulsory Schooling Reforms, Education and Mortality in Twentieth Century Europe. University of Wuppertal (mimeo).

Glied, S., Lleras-Muney, A., 2008. Health Inequality, Education and Medical Innovation. Demography 45 (3), 741–761.

Greene, W., 2003. Econometric Analysis. Prentice Hall, New Jersey.

Heck, K., Pamuk, E., 1997. Explaining the relation between education and postmenopausal breast cancer. American Journal of Epidemiology 145 (4), 366–372.

Hemminki, K., Li, X.J., 2003. Level of education and the risk of cancer in Sweden. Cancer Epidemiology Biomarkers Preview 12, 796–802.

Hemminki, K., Li, X.J., 2004. University and medical education and the risk of cancer in Sweden. European Journal of Cancer Prevention 13, 199–205.

Holmlund, H., 2007. A Researcher's Guide to the Swedish Compulsory School Reform. Working paper 9/2007, Swedish Institute for Social Research, Stockholm University.

Honoré, Bo E., Lleras-Muney, A., 2006. Bounds in competing risks models and the war on cancer. Econometrica 74 (6), 1675–1698.

Hussain, S.K., Altieri, A., Sundquist, J., Hemminki, K., 2008. Influence of education level on breast cancer risk and survival in Sweden between 1990 and 2004. International Journal of Cancer 122 (1), 165–169.

Lager, A., Torssander, J., 2012. Causal effect of education on mortality in a quasi-experiment on 1.2 million Swedes. Proceedings of the National Academy of Sciences 109 (22), 8461–8466.

Lagerlund, M., et al., 2002. Sociodemographic predictors of non-attendance at invitational mammography screening – a population-based register study. Cancer Causes and Control 13 (1), 73–82.

Lange, F., 2011. Education and allocative efficiency. Evidence from cancer screening. Journal of Health Economics 30, 43–54.

Lichtenberg, F., Lleras-Muney, A., 2005. The effect of education on medical technology adoption: are the more educated more likely to use new drugs? Special issue of the Annales d'Economie et Statistique in memory of Zvi Griliches, No 79/80.

Lleras-Muney, A., 2005. The relationship between education and adult mortality in the U.S. Review of Economic Studies 72 (1), 189–221.

Lund, E., Jacobsen, B.K., 1991. Education and breast-cancer mortality – experience from a Large Norwegian Cohort Study. Cancer Causes and Control 2, 235–238.

Marklund, S., 1981. Skolsverige 1950–1975: Försöksverksamheten. Liber Utbildningsförlaget, Stockholm.

Meghir, C., Palme, M., 2003. Ability Parental Background and Educational Policy: Empirical Evidence from a Social Experiment. Institute for Fiscal Studies, IFS Working Papers: W03/05.

Meghir, C., Palme, M., 2005. Educational reform, ability, and family background. American Economic Review 95 (1), 414–424.

Meghir, C., Palme, M., Schnabel, M., 2011. The Effect of Education Policy on Crime: An Intergenerational Perspective. NBER Working Paper 18145.

Meghir, C., Palme, M., Simeonova, E., 2012. Education, Health and Mortality: Evidence from a Social Experiment. NBER Working Paper 17932.

Meghir, C., Palme, M., Simeonova, E., 2013. Education, Cognition and Health: Evidence from a Social Experiment. NBER Working Paper 19002.
Nechuta, S., Paneth, N., Velie, E., 2010. Pregnancy characteristics and maternal breast cancer risk: a review of the epidemiologic literature. Cancer Causes and Control 21, 967–989.
Oreopoulos, P., 2006. Estimating average and local average treatment effects of education when compulsory school laws really matter. American Economic Review 96 (1), 152–175.
Spasojevic, J., 2010. Effects of education on adult health in Sweden: results from a natural experiment. Contributions to Economic Analysis 290, 179–199.
Storey, J., 2002. A direct approach to false discovery rates. Journal of the Royal Statistical Society: Series B 64, 479–498.
Storey, J., Tibshirani, R., 2003. Statistical significance for genome-wide experiments. Proceeding of the National Academy of Sciences 100, 9440–9445.
Tabar, L., et al., 1985. Reduction in mortality from breast cancer after mass screening with mammography. The Lancet 325 (8433), 829–832.