Sovereign risk and fertility: a natural experiment on Italy

Chiara L. Comolli

Abstract

Studies documenting the pro-cyclicality of fertility to business cycles in advanced economies mostly investigate the relationship between labor market indicators and birth rates. However, part of the recent fertility drop witnessed in Europe after the onset of the Great Recession is not explained by labor market measures. Using the case of the sovereign debt crisis of 2011 in Italy, this study shows that a significant drop in births rates was caused by perceived financial uncertainty. The spread between the Italian and German government bonds became a thermometer of the second phase of the Great Recession, measuring the skyrocketing interest rate in the Italian public debt. This sovereign debt crisis was caused by the loss of credibility from the financial markets on Italy being able to replay its debt. Perceived uncertainty is measured using Google trends searches for the term Spread to capture the degree of concern to general public about the stability of Italian public finances. The regression discontinuity in time research design identifies the causal effect of perceived uncertainty on monthly birth rates at the national and regional level in Italy. Results show that a drop of 4 births per 1000 women in reproductive age was caused by this concern, nine months later. A number of robustness checks corroborate these results.

Acknowledgements

The research leading to these results has received funding from the Strategic Research Council of the Academy of Finland (Decision Number: 293103) for the research consortium Tackling Inequality in Time of Austerity (TITA). I am also grateful for financial support from the Swedish Research Council (Vetenskapsrådet) via the Linnaeus Center for Social Policy and Family Dynamics in Europe (SPaDE), grant registration number 349-2007-8701. I thank the SUDA Colloquium participants for their valuable comments.

1. Introduction

"More seems to be at work, however, than these mechanical forces - namely, a general feeling of uncertainty. Assessing the precise nature and effects of this uncertainty is essential, but it is not easy. […] Uncertainty appears more diffuse, more *Knightian* in nature."
Olivier Blanchard[1]

Uncertainty is a widespread feature of contemporary societies. It represents a pervasive component of individual identities and social structures (Giddens 1991) and its consequences are vast (Halpern 2005; Keynes 1921; Knight [1921] 2002). In globalized societies, deregulation, internationalization and delocalization processes generate this permanent component of uncertainty (Blossfeld, Mills and Bernardi 2006; Blossfeld and Hofmeister 2006). In addition, macro and micro level uncertainty spikes when sudden shocks such as economic crisis, conflicts, natural disasters or social unrest produce unpredictability. In general terms, this lack of predictability, of clarity of future events constitute uncertainty and induce individuals to diversify their behavior (Johnson Hanks 2004, 2006). The concept of uncertainty, and in particular economic uncertainty, rapidly entered and spread among social scientists and their research, recently also among family demographers (Busetta, Mendola and Vignoli 2019). The discourse around uncertainty largely focus around three issues: the distinction of uncertainty from risk, the measurement of uncertainty and the proof of causal link between uncertainty and 'real quantities'.

Distinguishing it from risk (under which actors take decisions linked to a set of possible outcomes each one associated with a known probability) the American economist Frank H. Knight (1921) coined the current definition of uncertainty[2] as the condition under which, instead, actors' ability of assigning a probability distribution to future outcomes is impossible. According to this definition, the main distinction between risk and uncertainty is thus that the former is quantifiable; the latter is not.

Some authors, however, more recently argue that, despite being two different phenomena, both risk and uncertainty are measurable Dequech (1999, 2000, 2003). Due to the very broad formal definition of uncertainty, it is difficult to empirically separate it from risk and, in fact, most of the existing measures are a mixture of indicators of risk and uncertainty (Bloom 2014).

In economics, volatility measures are the most diffused proxy of aggregate economic uncertainty because the more volatile a series is the less it is predictable. Uncertainty indicators at the macro-level such as the volatility in GDP growth, stock market or professional forecasts all rise sharply

[1] Foreword of the IMF Outlook of October 2012 (pp. XV).
[2] In the quotation above, Olivier Blanchard refers to Knightian uncertainty talking about the Great Recession.

during the recessions. Another diffused indicator is the frequency of newspaper articles referring to economic uncertainty which also increases during economic crisis (Baker, Bloom and Davis 2012).
In sociology and demography economic uncertainty is measured very differently. At the aggregate level, macroeconomic indicators are sometimes used to measure the uncertain 'climate' individuals live in. Examples include consumer confidence, or labor market indicators such as long-term unemployment rates or the share of temporary or fixed-term job contract in a given year. At the individual level, uncertainty is measured through either direct questions to respondents about expectations or perceptions about the future situation (Kreyenfeld 2016), through other survey questions about their knowledge about possible events (Trinitapoli and Yeatman 2011) or individual indexes of joblessness persistence (Busetta, Mendola and Vignoli 2019).
Economics and demographic studies have led to important insights into how uncertainty is associated to reproductive strategies. Due to the irreversibility of childbearing decisions and the possibility of postponing birth for a certain time, uncertainty may generate a re-evaluations of preferences, risk and opportunities that manifest through either a permanent decline or a temporary postponement of childbearing (Ranjan 1999). Caldwell (2004) argues that social upheaval, by creating uncertainty about the future, can accelerate fertility declines related to the demographic transition. Other macro-level studies conducted on developed countries demonstrates that economic uncertainty lead to delays in childbearing (Adsera 2011; Mills and Blossfeld 2003). This has been illustrated in general in relation to the current globalization processes (Blossfeld, Buchholz, and Hofäcker 2006; Blossfeld and Hofmeister 2007; Blossfeld et al. 2009; Oppenheimer 2003) but also in relation to more specific events. For instance, the drop of birth rates in the aftermath of the Great Depression (Morgan 1991; Ryder 1980), or in Eastern Europe after the collapse of the communist system and the transition to a market economy (Ranjan 1999; Billingsley 2011; Sobotka et al. 2011). More recently, since the Great Recession hit advanced economies after 2007, economic uncertainty has often been cited in explanations of the most recent fertility rate decline in contemporary societies (Sobotka et al. 2011; Lanzieri 2013; Goldstein et al. 2013; Comolli 2017; Kreyenfeld 2016; Trinitapoli & Yeatman 2011).

All those studies, however, suffer from two main limitations. First, the operationalization of economic uncertainty is generally limited to traditional macroeconomic indicators that do not capture entirely the relationship between uncertainty and fertility. Second, the causal nexus between the latter is usually not addressed. This is the additional issue debated in the literature on uncertainty, namely its causal nexus with 'real quantities'. Economists are interested in the causal relationship between economic uncertainty and economic growth, unemployment rates or investment and consumption dynamics. Sociologists and demographers are interested in the causal effect of economic uncertainty on family decisions, among which childbearing is a primary example.

Far from identifying an omnibus measure of economic uncertainty, the first aim of the current study is to explore one alternative operationalization of uncertainty that is still ‘aggregate’ but less domain-specific than the macroeconomic indicators such as the share of the long-term unemployed or individuals with fixed-term job contract. The second objective is the identification of the causal effect of acute financial uncertainty on birth rates. In order to do that, the study is bounded to the case of the Sovereign debt crisis of 2011-2012 in Italy. In a later phase of the Great Recession, in 2011-12, Southern European countries have been hit by a confidence crisis on their public debts. The welfare provisions adopted during the early years of the crisis dramatically busted public expenditure and consequently government debts. Financial markets started doubting some states’ ability to ever replay their large and increasing debts, therefore started asking for a larger premium to buy those debts. The so-called Sovereign debt crisis materialized thus through the widening differentials in bond yields between risky, for instance, Italy or Greece, and safe countries, Germany. The financial speculative attack on public debts touched upon the very stability of those states in which a political, on top of the financial and economic, crisis exploded in 2011.

The identification strategy of the aggregate uncertainty produced by the Sovereign debt crisis, relies on the use of a new indicator, Google trends, to identify the tipping point in time of this uncertainty. The spread, namely the interest rate differential between Italy (risky) and Germany (solid) became the indicator of the financial credibility loss of the country, not just as a technical financial indicator of uncertainty but a crucial measure of how this uncertainty was perceived by the larger public. In the media narrative the term spread became the thermometer of the crisis, of ordinary use in the news (compared to other term like bailout, subprime or bankruptcy) and remained in the newspapers cover page for months. Many TV-shows were interviewing people and politicians in the streets asking if they knew what the word on everyone’s lips was. The problem with using the spread directly as the uncertainty indicator is that went up and down for months (IMF 2012; EBC), making it difficult to isolate a single shock and identify a causal effect. However, being the term so popular, it is likely that people started searching for information on the spread and many might look for information through internet. Based on this assumption, I use the Google searches for the word ‘spread’ as a proxy for when the interest in the topic peaked and isolate the uncertainty spike. Google Trends provides a normalized time series index corresponding to the volume of the queries of the users that were introduced into Google in a specific geographic zone (Choi and Varian, 2012). Google searches have been used before as a measure of issue salience, meaning that they “track general concern about the issue” (Mellon 2013: 280).

2. Background

2.1. Economic uncertainty and fertility

The pro-cyclicality to business cycles of fertility rates, increasing during periods of economic growth and decreasing during recession, has been extensively investigated after the onset of the Great Recession (Sobotka et al. 2011; Goldstein et al. 2013; Comolli 2017). Investigations based on aggregate-level fertility rates appear to support the idea that economic uncertainty is related to variation in national fertility rates (Goldstein et al. 2013; Lanzieri 2013; Sobotka et al. 2011). With few exceptions, scholars have demonstrated the negative association between various measures of labor market deterioration and birth rates. Rising unemployment rates (male, female or age-specific) are the first and principal suspect for the explanation of why childbearing is postponed during recessions. Unemployment rates are strongly and robustly correlated to TFR drops during, not only the most recent economic and financial crisis, but in all the main economic downturns of the last centuries in advanced economies (Sobotka et al. 2011). The same cannot be said of any other indicator used to explain the fertility decline such as GDP, housing market indicators or consumer confidence even though even the latter are all somehow measures of economic instability. However, all these existing indicators together do not explain entirely the substantial, prolonged and very diffused decline in birth rates in the European countries after the Great Recession. First, analyses that include a number of indicators such as unemployment rate, economic policy uncertainty index, the cost of public debt and consumer confidence index simultaneously do well but do not explain all the decline in birth rates in Europe and the US in the period 2000-2013 (Comolli 2017). Second, the drop in fertility took place even in the Scandinavian countries after 2011 and the latter is not explained by traditional macroeconomic indicators (Comolli 2018). Most puzzling are the trends in the Nordic countries, which have seen radical declines in their period fertility rates since 2010. In Denmark, for example, the total fertility rate dropped from 1.9 in 2010 to 1.7 in 2013; and for no apparent reason (Statistics Denmark 2014). This decline was no less radical than the decrease in fertility rates seen in Greece after its economy collapsed (Kreyenfeld 2016). Finally, a recent paper (Buckles, Hungerman & Lugauer 2018) shows that fertility decline in the US actually anticipated several recessions in the last decades. These few examples show that there is something that drives childbearing postponement during periods of economic crisis that is not captured by traditional indicators.

One argument usually found in the literature is that this unexplained share of fertility declines during economic downturns is due to perceived uncertainty, not so much linked to the actual economic outlook but associated to the perception, or anticipation, of future downturns (Trinitapoli

& Yeatman 2011; Kreyenfeld et al. 2012; Kreyenfeld 2016; Hofmann & Hohmeyer 2013). The problem is how to measure this perceived uncertainty, different from risk because uncertainty is what models cannot predict. In economics stock market volatility is often used but a recent paper (Jurado et al. 2015) shows that the number of relevant shocks is much lower than what stock market volatility would predict but these few shocks are so strong that they have much more persistent effects on the real economy. Furthermore, causality has not been proved so the question of whether uncertainty causes recessions or the other way around is still unanswered. In the demographic and sociological literature other measures like long-term unemployment, the share of self-employed (Matysiak et al. 2018) or the share of precarious or temporary contracts in the economy (Vignoli et al. 2012, 2016) are used. However, it is difficult to separate the latter from the previously mentioned macroeconomic traditional indicators because they are strongly correlated.

2.2. Google searches and issue salience

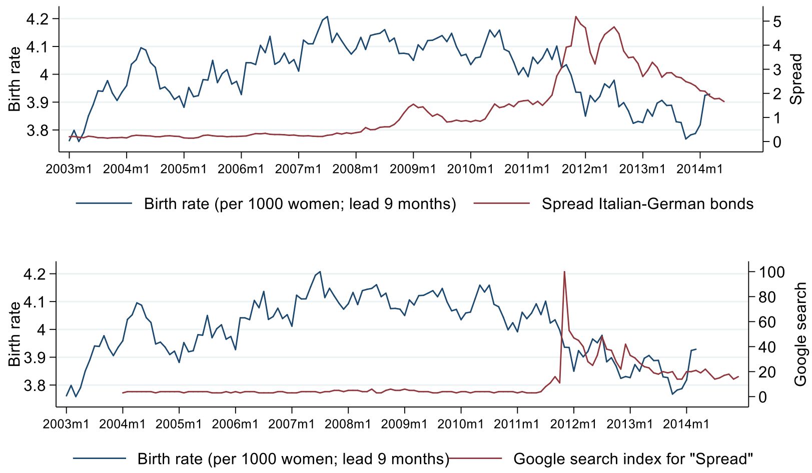
A small but growing body of research exploring the socioeconomic and demographic implications of growing internet diffusion and use. Part of this literature explores in particular the meaning of web searches and their implications for individuals' decisions. The main mechanisms linking the two is information gathering: broadband availability and web searches reduce the cost of seeking information with respect to the more decentralized offline information markets (Guldi and Herbst 2017). Between 2007 and 2017 print newspapers read declined from 67% to 35.8% of Italians. In the same period households' expenses on print newspapers and book declined 37% (CENSIS 2018). In 2014 the share of Italians reading online newspapers was 20.8% and reading news-related websites is 34.3% (CENSIS 2018).

The Google Trends tool was introduced in the summer of 2008 to provide a public view into the relative internet search volumes. Google Trends provides a time series index of the volume of the queries into Google in a specific geographic area (Choi and Varian, 2012). The advantage of Google search activity data is that they express the demand for a wide range of information and they are seen as involuntary, indirect data that can be more revealing than classical survey data. Search data allow to investigate combinations of space, time and context related to many dimensions of human behavior. Finally, the data are high in frequency and are available almost in real time. The drawbacks of Google search activity data are that they are available only in aggregate form and the methodology of collection and reporting of the data is not very transparent. Moreover, the geographical distribution of the searches is not always precisely estimated since the IPs cannot always be properly located. The meaning of some terms may change over time and places (Askitas 2015).

The use of search behavior as a proxy was popularized by Ginsberg et al. (2009) with Google Flu article. Google trends are mostly recognized as useful to measure the issue salience of a topic among the general public (Mellon 2013; 2015) and for the "tracking of real-life quantities" (Ojala et al 2017). Google trends have been used to track *economic indicators,* such as 'job search' for unemployment in various countries (Germany, Italy, Spain) (Choi and Varian 2012; Simionescu and Zimmermann 2017 for a review). Mellon (2013) shows that for Spain, even with low internet penetration levels, GT searches generate search data that closely match survey measures, especially for economic terms (Mellon 2013: 289). In Italy internet coverage of households is very similar to Spain (Fig. A.1-2). Naccarato et al. (2015) showed that Google search (e.g. query 'offerte di lavoro' meaning job offers) is a useful tool in nowcasting the Italian unemployment rate. D'Amuri and Marcucci (2009) and D'Amuri (2009) also find that 'Offerte di lavoro' is the most popular keyword used for job searches in Italy. Francesco (2009) shows that models using Google search data improve out-of sample prediction of unemployment rate in Italy. In relation to fertility, searches for the term 'maternity' have been shown useful to *forecast fertility* (Billari et al 2013; Ojala et al 2017) and the differences in tracking temporal vs. spatial variation in fertility (*Meaning of fertility* in different context, Ojala et al 2017). Finally, in relation to the consequences of the Great Recession google queries for "malaise" and "symptoms" have been used to track the effects on health (Askitas and Zimmermann 2015) and searches for "hardship letter" in the US have been used as an indicator of mortgage delinquency to track the effects of the crisis on the housing market (Askitas and Zimmermann 2011).

Figure 1 shows that google searches for the term spread peaked in Italy in November 2011. They were almost zero before that date and gradually returned to close to zero today. The date of the maximum searches can identify when the sovereign debt crisis in Italy was perceived as most salient among the population. The latter can thus be used to investigate the effects of that uncertainty on birth rates by looking at what happened 9 months after that peak.

Figure 1: Birth rates, Spread and Google queries for the term “spread”.



Source: Elaboration of the author based on Eurostat, ECB and Google data (2018).

3. Method

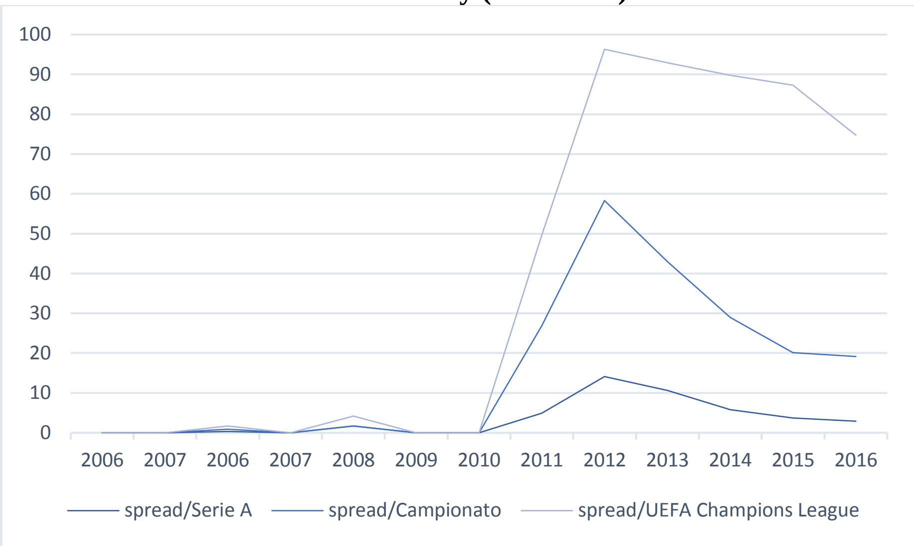
3.1. Data and measures

Birth rates have been derived from Istat (Istituto Nazionale di Statistica Italiano) dividing live births per month by the number of women 15-44 resident in Italy at January 1st each year. At the regional level, monthly live births in each region are divided by the number of women 15-44 resident in the region at January 1st each year. Birth rates (per 1000 women) are de-seasonalized using a centered 12-months moving average (6 months before and 6 months after each month). Finally, birth rates are lead 9 months to capture the conditions at the time of conception.

Google trends (GT) data (for a detailed non-technical description of how data are collected and released see Askitas 2015) represent the searches for the word "Spread' in www.google.it. The values represent search interest relative to the highest point on the chart for the given region and time [$\frac{\#Searches}{Maximum\#} * 100$]. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. A score of 0 means that there was not enough data for this term. Figure 2 shows the change in popularity of the term spread over the period 2006-2016 in Italy relative to searches regarding very common terms used in Italy to inquire about football: "Serie A" (A league), "Campionato" (Championship) and "UEFA Champions League". Until 2010 the queries for spread were null compared to any of the other three terms. In 2012 the queries for spread become almost 100 times those of UEFA Champions League, 60 times that of Campionato and 15 times that of Serie A and gradually decline in the following years.

The system eliminates repeated queries from a single user over a short period of time, so that the level of interest isn't artificially impacted by these type of queries (for instance, by typing errors). Google Trends at national level used in this analysis are robust to specific (related) search category: all categories, Finance, People and society, while as expected the trend values are different in non-related categories (e.g. food &drinks). Estimates are also robust depending on whether the search is done in Web search (general) or News search. Finally, complements, namely searches for the term 'spread' in non-related words (*Spread Eagle*, US band; *Spreadsheet*, Excel; *Spread shirt*, clothes printing online shop) have been removed. Figure 3 shows the importance of such robustness. A related event to the burst of the Sovereign debt crisis in Italy was the resignation of the government (Berlusconi was the Prime Minister at the time), which also took place in November 2011. Predictably, the search for the term "dimissioni Berlusconi" (resignation Berlusconi) peaked at the same time in the News category search on Google (panel a), however, the query was not so prevalent as that of 'spread' when we consider the general category (panel b).

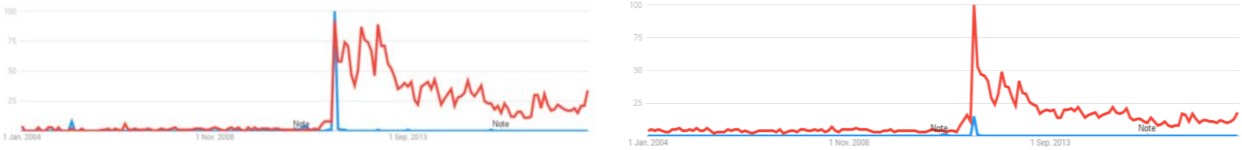
Figure 2: Ratios of Google Trends queries for Spread vs. UEFA champions league, Campionato and Serie A in Italy (2006-2016).



Source: Elaboration of the author based on Google Trends data (2018). Searches in the general All categories (annual average of monthly searches).

Figure 3: Google Trends for "spread" (red line) and "dimissioni Berlusconi" (blue line)

News category (a) | General category (b)



Source: Elaboration of the author based on Google Trends data (2018).

Geo-localization allows to see in more specific locations how popular the query for 'spread' is during the specified time frame. The value of 100 is given to the date with the greatest popularity as a fraction of total searches in that location in the chosen period. As in case of country-level data a higher GT value means a higher proportion of all queries, not a higher absolute query count. A key limitation with GT at regional level is data sparsity, in fact, zeroes indicate a location where there was not enough data for queries on this term. When modelling temporal variation at the regional level, GT fails to provide any data in 926/2640 (35.1% of cases) because there were not enough queries on those region/month combinations (reported as missing data in the GT variable). In the regional data, queries in all categories and no complements have been considered.

The treatment variable is a dummy for every month after the peak in searches (GT=100). In the main analysis I do not use the search interest index itself, although it is added in some models for robustness check. It is worth noting that, although the treatment dummy in the model is allowed to differ by region (equal 1 when GT=100) in almost all regions for the period 2004-2018 the peak in GT was registered in November 2011. The only two exceptions are Val d'Aosta where the peak in GT queries for 'spread' took place in December 2011 and Molise that registered the peak in March 2012.

Finally, other variables added in models for robustness checks are monthly marriage rates, the actual spread between Italian and German sovereign bonds, GT index, unemployment rate, consumer confidence index and some of the more detailed indicators used to construct it (confidence in own and general economic situation, confidence in current and future economic situation).

3.2. Research design

The gold standard for measuring the impact of an event (or policy) the randomized controlled experiment in which the units exposed to the treatment are the same in expectation to those who were not, the control group. It is clear that large-scale experiments allocating economic uncertainty to some individuals only, or recessions to some regions only, are not ethically feasible. In the lack of such randomization process we need to rely on observational data to determine the impact of the exposure to such events. However, non-experimental data pose many challenges to the distinction of correlation from causality and the identification of the latter (Angrist and Krueger 2000). Regression discontinuity (RDD) represents one strategy to identify the causal effect of a treatment using observational data in which the treatment is based on a 'forcing' variable, such as location, birthdate or time, being below or above a certain threshold. The identifying assumption is that there is no discrete change in the characteristics of individuals just below or above the threshold for treatment assignment (Athey and Imbens 2016). Units around the cutoff, therefore, do not systematically differ in their unobservable characteristics, offering a valuable counterfactual comparison between control and treatment group (Calonico et al 2016). A key feature of RDD is simplicity and transparency.

When as in the present case time is the forcing variable, the design is a particular case of RDD and it is called Regression Discontinuity in Time (RDiT). At the time of writing RDiT is mostly used in the fields of public economics, industrial organization, environmental economics, marketing and international trade to estimate the treatment effect of policies (Hausman and Rapson 2017)[3]. What is relevant here is that RDiT design is often used in similar contexts to, but it is not identical to, first,

[3] Most papers using RDiT are in environmental economics: Anderson (2014); Auffhammer and Kellogg (2011); Bento et al. (2014); Burger et al. (2014); Davis (2008); Davis and Kahn (2010); Chen et al. (2009); Chen and Whalley (2012); De Paola et al. (2013); Gallego et al. (2013); Grainger and Costello (2014).

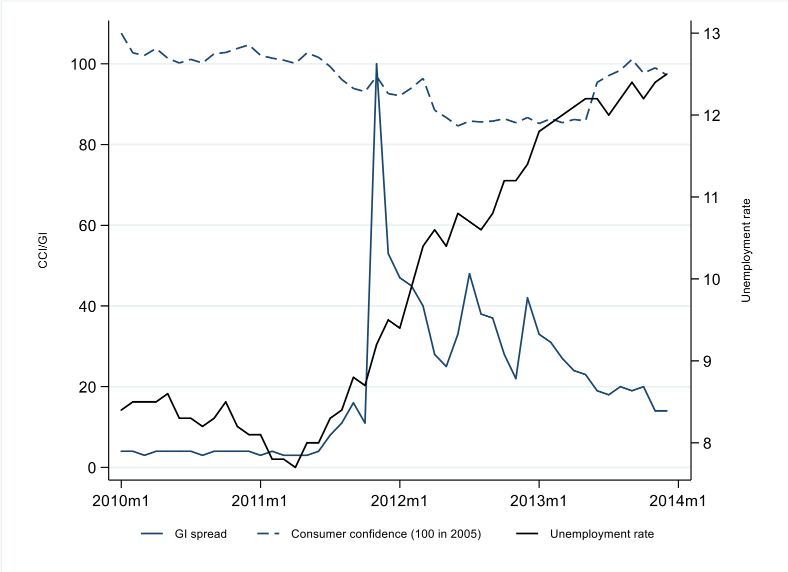
cross-sectional RDD (Shadish et al., 2002) and, second, to event study or interrupted time series regression methods (Bernal, Cummins and Gasparrini 2017).

RDiT differ from event studies or interrupted time series because the latter, first, dispose usually of high-frequency data and, consequently, focus on short-time windows (T) around the event, second, dispose of many cross-sectional units (N) and, third, do not use high-order polynomial controls in time (Hausman and Rapson 2017). One advantage to use RDiT is that, in contrast to event studies it is not necessary to assume that there are no unobservable variables correlated with time, it is enough to assume – with the caveats expressed below – that the latter do not change discontinuously at the threshold (Davis 2008; Hausman and Rapson 2017).

The cross-sectional vs time-series type of data represent the first obvious difference between RDD and RDiT. In cross-sectional RDD one needs to have a large-enough sample (N) in the neighbourhood of the cut-off to approximate the limit of the conditional expectation of the outcome variable just-below and just-above, however, in an RDiT there is little or no cross-sectional variation. This could be problematic in light of the well-known bias/precision trade-off as the sample size increases away from the cut-off by increasing T instead of N as in standard RDD analyses. Researchers must rely on observations away from the threshold in order to obtain sufficient power to precisely estimate the coefficient and using observation remote in time from the cut-off "represents a substantial conceptual departure from the identifying assumption used a cross-sectional RD" (Hausman and Rapson 2017: pp.3). Therefore, if unobservable confounders (for instance: self-selection into treatment, anticipation behaviour) or time-series properties are not correctly addressed, estimates could be biased. Assuming continuity of unobservables at the threshold, normally enough with standard RDD to ensure identification, is not enough in the case of RDiT, for which there are three additional issues: anticipation close to the threshold, time-series auto-regression and short- vs long-term effects of the treatment. Recommendations from the literature propose presenting the following robustness checks: global polynomial and local linear; various order of polynomial; placebo tests using other dates or nearby areas not affected by the treatment; plot RD estimates on control variables to show continuity. All of them are included in the present study.

The identification assumption at the basis of this paper is thus that births that take place around 9 months after the peak in spread (just before or just after August 2012) do not differ on parents' unobservable characteristics, the only unsmooth change at the cutoff is due to the sudden surge in the spread salience, namely in uncertainty. Figure 4 graphically shows that on the treatment date, in November 2011, the jump is due to the treatment only while other determinants of fertility are smooth on the threshold. Figure 4 shows the Google Trends for spread, the unemployment rate and the consumer confidence index.

Figure 4: Trends in unemployment rate, consumer confidence and Google queries for spread.



Source: Elaboration of the author based on Google Trends, Eurostat and Istat data (2018).

As mentioned, RDiT is affected by the tradeoff between precision and bias. The researcher wants to stay as close as possible to the cutoff (where observations are more similar) but you want to have enough data points to get precise estimates. The assumed functional form of the relation also determines how close one can stay to the cutoff: the closer to the cutoff the more linear the relationship gets. If the underlying regression function of Y is fairly linear, the bandwidth can be larger to get more precise estimates without loss in terms of bias; however, if the regression line of Y is non-linear, the bandwidth should be restricted to get unbiased estimates. Since there is no prior knowledge to assume that the functional form of the births rates might be linear (Local linear regression, Eq. 1) rather than polynomial of any degree (Global Polynomial estimation, Eq. 2), the underlying regression function is estimated using both approaches. Moreover, the Optimal data driven algorithm derived by Calonico et al. (2014, 2015) is also presented. All three approached lead to very close estimates of the jump in birth rates nine months after the peak in uncertainty.

$$MA(\,Birth\,rate_{t+9}) = \gamma_0 + \gamma_1 D + \gamma_2 t + \gamma_3 D * t + \varepsilon_t \qquad \text{with } -h < T <= +h \qquad (1)$$

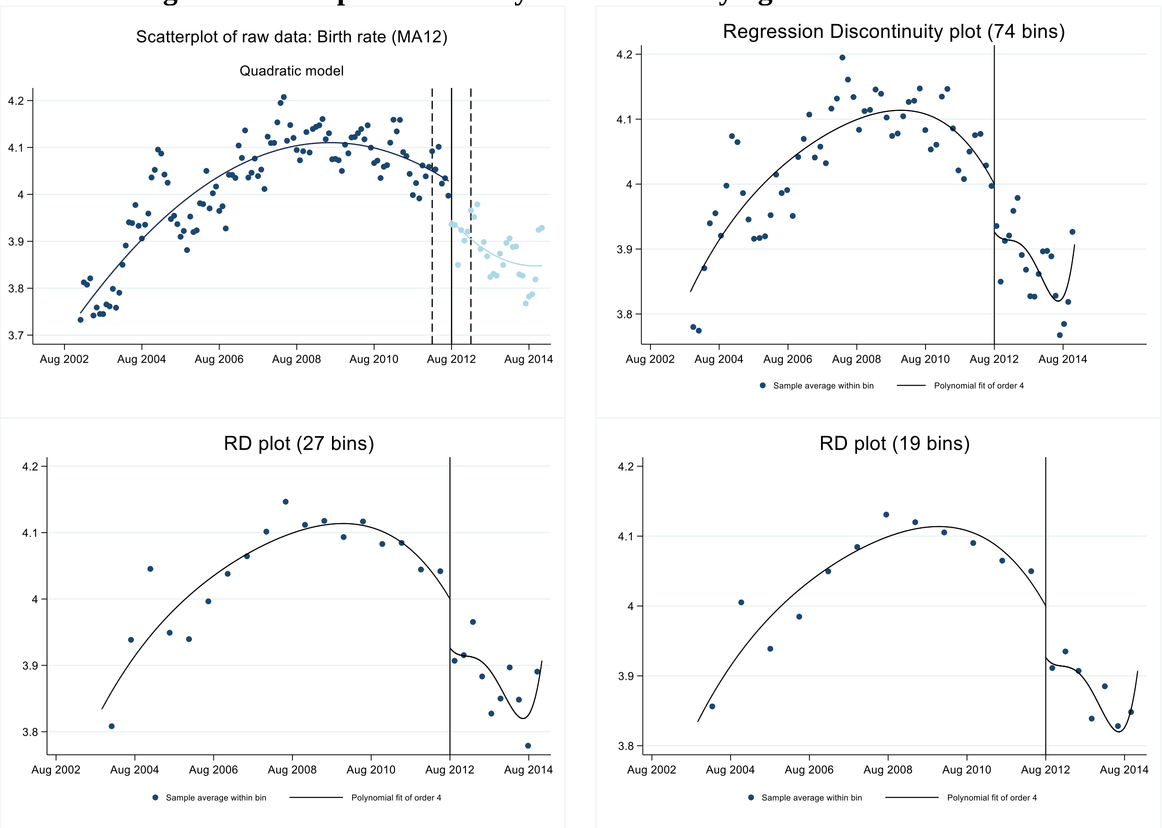
$$MA(\,Birth\,rate_{t+9}) = \gamma_0 + \gamma_1 D + \gamma_2 t + \gamma_3 D * t + \textstyle\sum_{k=2}^{P} \delta_k\, t^k + \sum_{k=2}^{P} \beta_k D * t^k + \varepsilon_t \qquad (2)$$

On the LHS we have the 12 months moving average birth rate at time *t* (leaded 9 months), and on the RHS we have the continuous time variable *t (year-month)*, centered around the cutoff date (0 in November 2011). *D* is the treatment dummy variable for after November 2011 (D=1 if t ≥ November 2011 and D=0 if t < November 2011), indicating the eventual departure from the trend occurring after the peak in Google searches for spread. Parameters are allowed to differ on the left (γ_2) and on the right (γ_3) of the cutoff (Nov 2011), because if we were to constraint the slope to be identical on both sides of the cutoff, the would amount to using data on the RHS of the cutoff estimate the effect, which is inconsistent with the spirit of RD (RHD is treated units, LHS is control). The main coefficient of interest, the jump at the cutoff is γ_1. The last polynomial term in Eq. 2 is added for the time trend for global polynomial estimation. Moving Averages are autoregressive, so to account for the dependence in the residuals, I clustered standard errors for years. Additional checks have been conducted using a AR(1) process. Local linear regression results are presented using different bandwidths (as mentioned, to balance precision and bias) and global polynomial are presented for different polynomial orders. If the underlying regression function of Y is fairly linear, the optimal bandwidth would be larger to get more precise estimates without loss in terms of bias; but if the regression line of Y is non-linear, the optimal bandwidth would be restricted to get unbiased estimates. A rectangular kernel - observations are not weighted depending on their distance from the cutoff - is preferred to a triangular kernel function as the former approach is more transparent (Lee and Lemieux 2010).

4. Results

Table A.1 in the Appendix reports summary statistics of the variables used in the analyses before and after the cutoff at the national and regional level. It shows that after the peak in uncertainty in November 2011, the financial and economic outlook of the country worsen together with the perception of economic conditions by consumers and the socio-demographic indicators. Both average crude birth and marriage rates in fact decline, all the consumer confidence indicators worsen, the actual spread on Italian bonds quadruple and the unemployment rate significantly increases. Figure 5 illustrates the RDiT plot of monthly birth rates for Italy (national estimates). The top-left panel shows the scatterplot of the raw data where each dot represents the birth rate of one month (N=144). The cutoff (leaded 9 months) point is in August 2012 and the vertical dotted line indicate an example of a 6 months badwidth. The other three panel report the same plot varying the width in which each bin is calculated averaging the birth rate over a given number of months (2 in the top-right, 5 in the bottom-left and 7 in the bottom-right panel). The figure illustrates the trade-off between precision and unbiasedness of the estimate of the jump at the threshold.

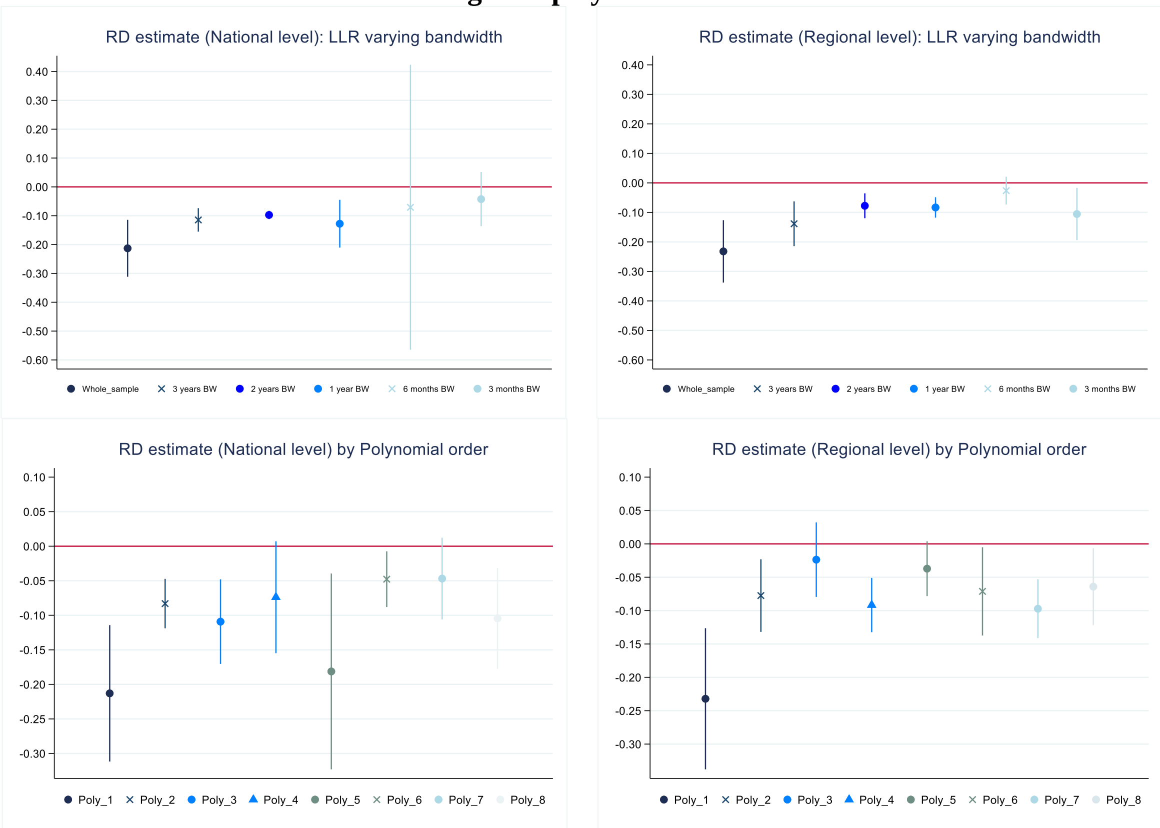
Figure 5: RDiT plot of monthly birth rates varying the number of bins.



Source: elboration of the author based on Istat data.

Figure 6 reports results of national (left panels) and regional (right panels) estimates for the local linear (top panel) and the global polynomial (bottom panels) RDiT estimates. Full models' estimates are reported in Tables A.2-5 in the Appendix. For the local linear, different estimates are reported varying the bandwidth: from using the whole sample or a symmetric bandwidth between 3 years and 3 months around the cutoff. For the global polynomial different estimates are presented varying the degree of the polynomial between the linear and an 8^{th} order polynomial. All estimates range between -0.20 and -0.10 with more precise estimates coming from the models using regional data as more observations are available. Additional estimates (not show, available from the authors upon request) simultaneously varying the order of the polynomial and reducing the bandwidth display similar coefficients. Taking a conservative estimate from the models, a drop of -0.10 in birth rate 9 months after the peak in GI for spread, considering the average rate of 4 births per 1000 women age 15-44 per month before the uncertainty burst, translates into around 2.5% drop in births due to the Sovereign Debt crisis. In size, this is similar to the associational evidence found between unemployment and Total Fertility Rates (-3%) and higher than the association between decline in consumer confidence and TFR (1%) for Europe and the US (Comolli 2017). The RDiT results are also stronger than the associational evidence in a region fixed effect model (not shown) of the correlation between GT searches for the term spread and monthly birth rates per 1000 women: -0.0034 (with a confidence interval of between -0.004 and-0.003, and intercept at 4 births per 1000 women).

Figure 6: RDiT coefficients plot. National and regional data. Local linear regression and varying order global polynomial.



Source: elboration of the author based on Istat data.

5. Robustness checks

5.1. Optimal data-driven RDiT

This section illustrates a number of robustness checks that confirm the results reported in the main analysis. Table 1-2 report results from the optimal data-driven calculation of the asymmetric bandwidth in local linear and different orders of polynomial RDiT. Although the estimates now at the regional-level do not reach statistical significance at the 95% confidence level, the magnitude of the coefficients are very close to those obtained in the main analyses.

Table 1: Optimal data-driven regression discontinuity estimates. National monthly data.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
RDiT estimate	-0.104***	-0.104**	-0.077	-0.089*	-0.086
Robust 95% CI	[-0.165 -0.044]	[-0.075 -0.033]	[-0.155 0.023]	[-0.0180 0.0003]	[-0.193 0.019]
N	135	135	135	135	135
Kernel Type	Uniform	Uniform	Uniform	Uniform	Uniform
Robust p-value	0.001	0.004	0.144	0.051	0.106
Order local polynomial (p)	1	2	3	4	5
Order bias (q)	2	3	4	5	6
BW-left	19.14	23.27	26.55	39.00	44.67
BW-right	6.63	9.31	9.40	13.06	13.51
N-left	19	23	26	39	44
N-right	7	10	10	14	14

Source: Istat and Google data (2017). Note: *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Optimal data-driven regression discontinuity estimates. Regional monthly data.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
RDiT estimate	-0.082	-0.093	-0.120	-0.095	-0.090
Robust 95% CI	[-0.217 0.045]	[-0.244 0.043]	[-0.300 0.041]	[-0.286 0.116]	[-0.303 0.130]
N	2631	2631	2631	2631	2631
Kernel Type	Uniform	Uniform	Uniform	Uniform	Uniform
Robust p-value	0.199	0.17	0.138	0.407	0.434
Order local polynomial (p)	1	2	3	4	5
Order bias (q)	2	3	4	5	6
BW-left	18.73	32.85	35.86	37.71	47.19
BW-right	8.33	13.7	17.68	19.38	20.72
N-left	360	640	700	740	940
N-right	180	280	360	400	420

Source: Istat and Google data (2017). Note: *** p<0.01, ** p<0.05, * p<0.1.

5.2. Confounding factors

Table 3 reports estimates obtained from the national sample, second order polynomial RDiT model using the entire period, controlling in the models for potential confounding factors, such as GT searches for the term spread, the Spread itself, unemployment rate and Consumer Confidence Index (CCI). The Table shows that the treatment effect remains substantially around -0.10.

Table 3: National-level 2nd order polynomial RDiT estimates, controlling for confounding variables.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
RDiT	-0.109***	-0.082***	-0.132***	-0.133***	-0.107***
	(-0.150 - -0.068)	(-0.109 - -0.055)	(-0.203 - -0.060)	(-0.224 - -0.042)	(-0.169 - -0.045)
Date (mean cent.)	-0.004***	-0.003	-0.004**	-0.004**	-0.004***
	(-0.006 - -0.002)	(-0.007 - 0.001)	(-0.007 - -0.001)	(-0.007 - -0.001)	(-0.007 - -0.002)
RDiT *Date	0.002	0.001	0.002	0.002	0.002
	(-0.002 - 0.007)	(-0.001 - 0.003)	(-0.002 - 0.007)	(-0.004 - 0.009)	(-0.003 - 0.008)
Date squared	-0.000***	-0.000**	-0.000***	-0.000**	-0.000***
	(-0.000 - -0.000)	(-0.000 - -0.000)	(-0.000 - -0.000)	(-0.000 - -0.000)	(-0.000 - -0.000)
Unemployment rate		-0.015			
		(-0.048 - 0.018)			
CCI			-0.001		
			(-0.004 - 0.001)		
GT spread				0.000	
				(-0.001 - 0.002)	
Spread					-0.001
					(-0.025 - 0.023)
Constant	4.028***	4.163***	4.157***	4.034***	4.030***
	(4.002 - 4.054)	(3.851 - 4.476)	(3.888 - 4.427)	(3.993 - 4.074)	(3.961 - 4.099)
N	135	132	123	123	135
R-squared	0.766	0.781	0.745	0.743	0.766

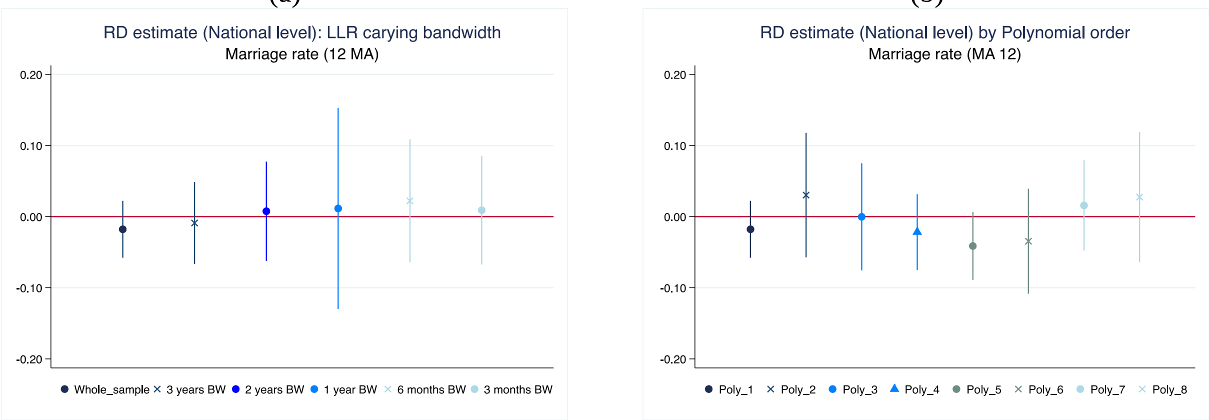
Source: Istat and Google data (2017). Note: *** p<0.01, ** p<0.05, * p<0.1.

5.3. Placebo

The last and most important robustness check tests a placebo treatment on monthly crude marriage rate nine months after the peak in uncertainty. While marriage rates are likely to be negatively influenced by economic uncertainty, due to a similar argument to childbearing (although marriages are reversible, at a certain cost) there is no reason to expect that this negative effect of uncertainty materializes nine months after the uncertainty shock, as it is the case for childbearing. Figure 7 illustrates both LLR and Polynomial RDiT estimates on marriage rates in Italy using national level data, varying respectively the bandwidth and the polynomial order. Monthly marriage rates at the regional level were not available at the time of writing so a similar placebo test using regional data could not be performed. The Figure shows that, as expected, the uncertainty shock had not effect on the national birth rates.

Figure 7: RDiT LLP and polynomial coefficients plot. National-level data.

(a) (b)



Source: elboration of the author based on Istat data.

6. Discussion

The emergence of new forms of risk is a consequence of the rising levels of complexity in modern industrial societies (Chappe 2012). With the globalization process, liberalizations and labor market deregulation, the probability of experiencing phenomena such as income drops, financial losses, unemployment or job precariousness, and downward mobility have increased. These new forms of risk emerged in contemporary societies and profoundly changed individuals and family behavior. Beyond this everyday risk component, some authors argue that the risk framework is not enough to capture the degree of unpredictability of events in contemporary society (Knight 1921). Uncertainty is thus defined as a condition of ignorance in which actors cannot predict the likelihood of the outcome produced by their actions. Olivier Blanchard refers to this Knightian definition of uncertainty in his foreword to the International Monetary Fund (IMF) Outlook in 2012, talking about the Great Recession and its consequences.

In this paper I investigate the causal nexus between economic uncertainty and fertility by studying the case of Italy in the aftermath of the Sovereign Debt crisis of 2011. In this second phase of the Great Recession, some countries like Italy suffered a loss of credibility in the financial market due to their skyrocketing public debts. The speculation on the inability of the country to repay its debt and the subsequent rise in the cost of the Italian debt was so sudden and brutal that the financial crisis rapidly escalated into a political crisis in which the very participation of Italy and other southern European countries to the EU was endangered. This uncertainty shock produced consequences way outside the public finance realm and individuals were affected by the insecurity produced. As Nau, Dwyer and Hodson (2017) argue: "there are many implications of having moved to a debt society that we are only beginning to understand […] young adults today must borrow against future in the hope that his investment pays off" (pp. 13). The spread between the cost of the risky Italian and the safe German bonds became a thermometer of this uncertainty in the media narrative and in everyday conversations. In a situation of ignorance, individuals seek information where they can and Google is one of these sources. The search queries for the term 'spread' suddenly spike in November 2011 when, I argue, the salience of the issue also peaked. I use this discontinuity to assess what have been the consequences on birth rates in Italy nine months after the peak.

Using monthly birth data from the Italian Institute of Statistics (ISTAT) at the national and at the regional level, I show that around 2.5% of the drop in births nine months after the uncertainty shock in November 2011 was due to the Sovereign Debt crisis. This result comes from the conservative estimate from the models, but considering the average rate of 4 births per 1000 women age 15-44 per month before the uncertainty burst, some point estimates suggest a drop in birth rate around -0.2 which translates into around a 5% drop in births due to the Sovereign Debt crisis. In size, this is similar or higher than the associational evidence between rising unemployment rate and declining Total Fertility Rates (-3%) and higher than the association between decline in consumer confidence and TFR (1%) for Europe and the US (Comolli 2017). Results are strongly corroborated by a few robustness checks including a placebo on marriage rates nine months after the shock, which as expected are not affected by the uncertainty spike.

The study suffers from a few limitations. First, as many research designs aimed at identifying the causal effect of events, RDiT favors internal validity at the expenses of external validity. In this case the question to ask is whether the causal effect identified here is very local, or maybe too local. Is a 2.5% drop the causal effect of economic uncertainty on fertility rates? I would not claim so given the particular case in which the uncertainty shock materialized. Second, it is difficult to say if what it is measured here is a consequence of a tipping point in perceived uncertainty or a consequence of Italians loosing money on their investment in state-debt obligations. One alternative mechanism, in fact, that could explain these results is that Italians anticipated financial losses in case of government default, which would imply that the eventual material wealth drop would explain the postponement of childbearing. On the one hand, as government bond are traditionally viewed as the lowest-risk financial investment possible, a drastic rise in their risk would induce private investors to sell their assets. This seems to be the case since after 2011 Italians drastically reduced their purchases of public debt. The share of public debt owned by Italian private citizens was in fact 21% in 2011 and 6% in 2018. On the other hand, however, debt is risky, it is a useful resource that can finance life transitions but it is also a liability that must be repaid. Previous studies show that the meaning of debt goes beyond its monetary component. Disposing of credit is seen as a way to security and affluence but an excess of debt is problematic and stressful (Nau, Dwyer and Hodson 2017). It is thus difficult to separate, in general, not only in the case of this paper, the strictly monetary component of debt and its value as an indicator of perceived uncertainty.

Third, despite the growing evidence of the usefulness of internet search data, the amount of studies using them did not reach a critical mass to be able to know exactly what individuals truly mean when they use web searches. There is a considerable amount of studies confident in the use of internet search data at least in the economic realm to predict macroeconomic indicators, but the use of them as a measure of economic uncertainty is a novelty and more research along the line of the current paper is needed.

References

Askitas, N. (2015). Google search activity data and breaking trends. *IZA World of Labor*.

Askitas, N., & Zimmermann, K. F. (2009). Google econometrics and unemployment forecasting. *Applied Economics Quarterly*, *55*(2), 107-120.

Askitas, N., & Zimmermann, K. F. (2015). Health and well-being in the great recession. *International Journal of Manpower*, *36*(1), 26-47.

Askitas, N., and K. F. Zimmermann. *Detecting Mortgage Delinquencies*. IZA Discussion Paper No. 5895, 2011.

Billari, F., D'Amuri, F., & Marcucci, J. (2016, October). Forecasting births using Google. In *CARMA 2016: 1st International Conference on Advanced Research Methods in Analytics* (pp. 119-119). Editorial Universitat Politècnica de València.

Buckles, K., Hungerman, D., & Lugauer, S. (2018). *Is Fertility a Leading Economic Indicator?* (No. w24355). National Bureau of Economic Research.

CENSIS (20128), 52° Rapporto sulla situazione sociale del Paese, Franco Angeli, Italia.

Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, *88*(s1), 2-9.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*(7232), 1012.

Guldi, M., & Herbst, C. M. (2017). Offline effects of online connecting: the impact of broadband diffusion on teen fertility decisions. *Journal of Population Economics*, *30*(1), 69-91.

Hofmann, B., & Hohmeyer, K. (2013). Perceived economic uncertainty and fertility: Evidence from a labor market reform. *Journal of Marriage and Family*, *75*(2), 503-521.

Jurado, K., Ludvigson, S. C., & Ng, S. (2015). Measuring uncertainty. *American Economic Review*, *105*(3), 1177-1216.

Kearney, M. S., & Levine, P. B. (2015). Media influences on social outcomes: The impact of MTV's 16 and pregnant on teen childbearing. *American Economic Review*, *105*(12), 3597-3632.

Knight, F. H. (1921). *Risk, Uncertainty, and Profit*. New York: Cosimo Classics.

Kreyenfeld, M. (2016). Economic uncertainty and fertility. In *Social Demography Forschung an der Schnittstelle von Soziologie und Demografie* (pp. 59-80). Springer VS, Wiesbaden.

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google Flu: traps in big data analysis. *Science*, *343*(6176), 1203-1205.

Matysiak, A., Sobotka, T. O. M. Á. Š., & Vignoli, D. (2018). *The Great Recession and Fertility in Europe: A Sub-National Analysis* (No. 1605). Vienna Institute of Demography (VID) of the Austrian Academy of Sciences in Vienna.

Mellon, J. (2013). Where and when can we use Google Trends to measure issue salience?. *PS: Political Science & Politics*, *46*(2), 280-290.

Mellon, J. (2014). Internet search data and issue salience: The properties of Google Trends as a measure of issue salience. *Journal of Elections, Public Opinion & Parties*, *24*(1), 45-72.

Nau, M., Dwyer, R. E., & Hodson, R. (2015). Can't afford a baby? Debt and young Americans. *Research in Social Stratification and Mobility*, *42*, 114-122.

Ojala, J., Zagheni, E., Billari, F. C., & Weber, I. (2017). Fertility and its meaning: Evidence from search behavior. *ICWSM'17*.

Simionescu, M., & Zimmermann, K. F. (2017). *Big Data and Unemployment Analysis* (No. 81). GLO Discussion Paper.

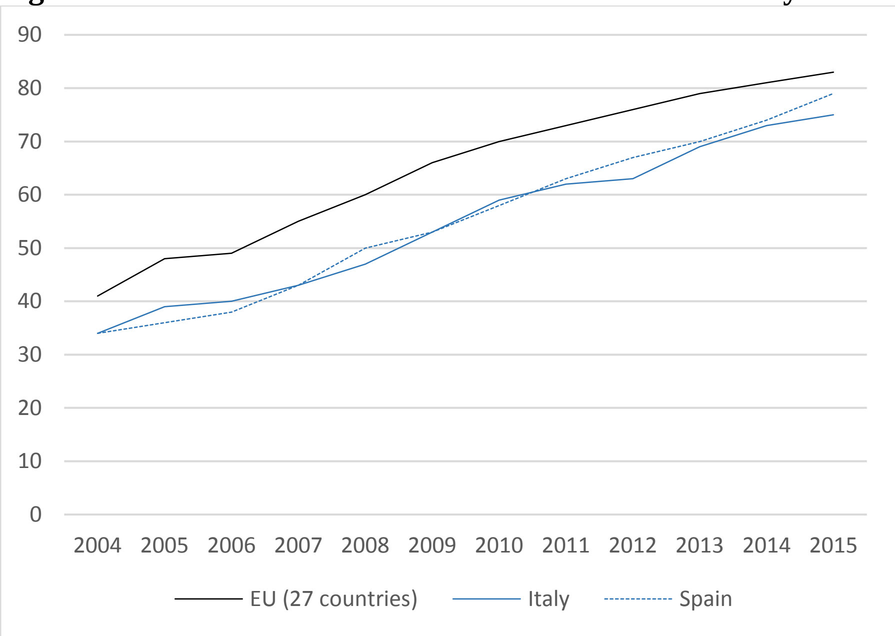
Trinitapoli, J., & Yeatman, S. (2011). Uncertainty and fertility in a generalized AIDS epidemic. *American Sociological Review*, *76*(6), 935-954.

Vignoli, D., Drefahl, S., and De Santis, G. (2012). Whose job instability affects the likelihood of becoming a parent in Italy? A tale of two partners. *Demographic Research* 26(2): 41-62.

Vignoli, D., Tocchioni, V., and Salvini, S. (2016). Uncertain lives: Insights into the role of job precariousness in union formation in Italy. *Demographic Research* 35: 253−282.

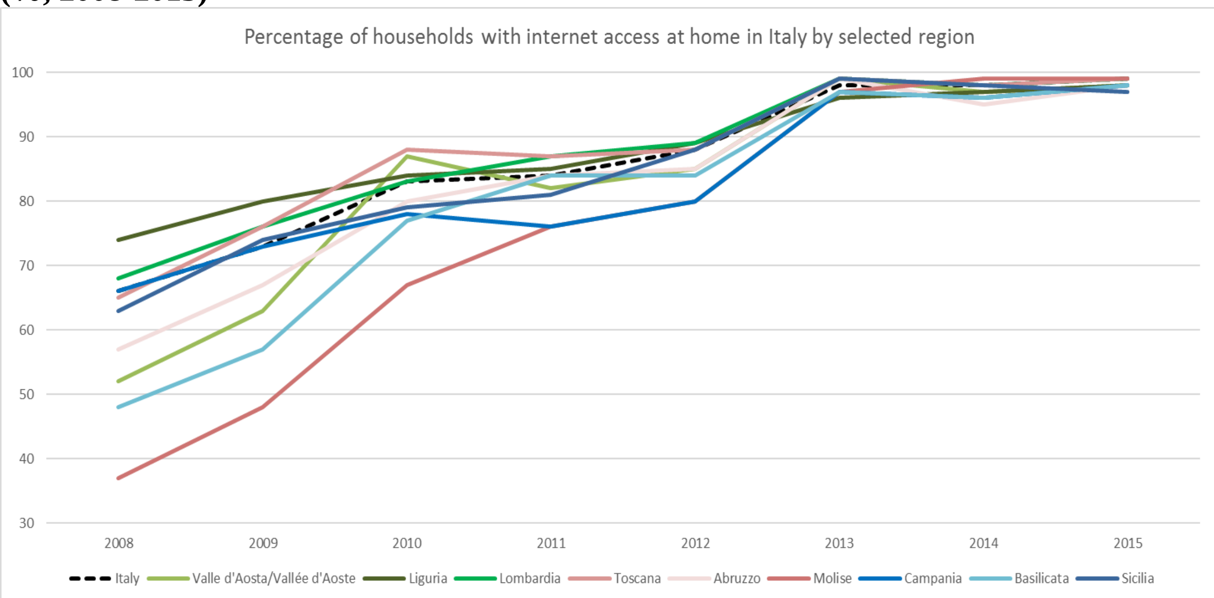
Appendix

Figure A.1: Households with internet access at home: Italy and EU average (%, 2004-2015)



Source: Eurostat (2017).

Figure A.2: Internet penetration by region in Italy ().Households with internet access at home (%, 2008-2015)



Source: Eurostat (2017).

Table A.1: Summary statistics

	BEFORE November 2011					AFTER November 2011				
NATIONAL ESTIMATES	**N**	**Mean**	**Std. Dev.**	**Min**	**Max**	**N**	**Mean**	**Std. Dev.**	**Min**	**Max**
Birth rate*1000	106	*4.03*	0.30	3.33	4.70	29	*3.90*	0.30	3.36	4.61
Marriage rate*1000	106	*0.33*	0.22	0.06	0.79	29	*0.27*	0.19	0.06	0.67
GT	94	*4.20*	1.81	3	16	38	*27.95*	16.13	14	100
Spread	106	*0.70*	0.75	0.14	3.97	32	*3.20*	1.02	1.66	5.19
Unemployment rate	106	*7.55*	0.82	5.80	8.80	26	*11.23*	1.02	9.20	12.50
Consumer confidence	94	*100.94*	3.81	92.20	108.90	38	*94.39*	7.16	84.60	105.90
Confidence on own situation (2005=100)	94	*101.47*	4.08	92.10	110.30	38	*96.85*	4.06	89.30	102.40
Confidence on general situation (2005=100)	94	*99.48*	10.25	73.00	121.80	38	*87.86*	16.73	58.10	116.90
Confidence on current situation (2005=100)	94	*103.60*	5.22	91.30	112.60	38	*97.10*	4.55	89.20	104.60
Confidence on future situation (2005=100)	94	*97.42*	5.32	82.10	107.20	38	*90.81*	12.22	69.60	108.90
	BEFORE November 2011					**AFTER November 2011**				
REGIONAL ESTIMATES	**N**	**Mean**	**Std. Dev.**	**Min**	**Max**	**N**	**Mean**	**Std. Dev.**	**Min**	**Max**
Birth rate*1000	1,880	*3.94*	0.47	2.65	6.09	751	*3.82*	0.51	2.54	5.94
GT	981	*6.87*	6.07	1	67	733	*28.90*	17.83	3	100
Unemployment rate	1,880	*9.77*	4.94	1.91	25.15	760	*9.68*	4.94	2.25	24.60

Table A.2: National-level LLR RDiT estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Whole_sample	Symmetric3_sample	Symmetric2_sample	Symmetric1_sample	Symmetric6m_sample	Symmetric3m_sample
VARIABLES	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag
dummy_or = 1	-0.213***	-0.115***	-0.097***	-0.128**	-0.071	-0.042
	(-0.312 - -0.114)	(-0.155 - -0.074)	(-0.112 - -0.083)	(-0.211 - -0.045)	(-0.565 - 0.423)	(-0.136 - 0.051)
date_c	0.002**	-0.002***	-0.002**	0.002	-0.017	-0.013
	(0.000 - 0.003)	(-0.003 - -0.001)	(-0.004 - -0.001)	(-0.009 - 0.013)	(-0.017 - -0.017)	(-0.013 - -0.013)
0b.dummy_or#co.date_c	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000 - 0.000)	(0.000 - 0.000)	(0.000 - 0.000)	(0.000 - 0.000)	(0.000 - 0.000)	(0.000 - 0.000)
1.dummy_or#c.date_c	-0.005***	-0.001	-0.002**	-0.003	0.014	-0.030
	(-0.009 - -0.002)	(-0.005 - 0.003)	(-0.004 - -0.000)	(-0.013 - 0.006)	(-0.110 - 0.139)	(-0.312 - 0.251)
Constant	4.139***	4.041***	4.034***	4.053***	3.989***	3.992
	(4.049 - 4.229)	(4.022 - 4.059)	(4.020 - 4.048)	(4.007 - 4.099)	(3.989 - 3.989)	(3.992 - 3.992)
Observations	135	65	48	24	12	6
R-squared	0.609	0.850	0.840	0.750	0.888	0.935

Table A.3: Regional-level LLR RDiT estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Whole_sample	Symmetric3_sample	Symmetric2_sample	Symmetric1_sample	Symmetric6m_sample	Symmetric3m_sample
VARIABLES	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag
dummy_or = 1	-0.232***	-0.138***	-0.078***	-0.083***	-0.026	-0.106**
	(-0.338 - -0.126)	(-0.214 - -0.063)	(-0.120 - -0.035)	(-0.118 - -0.049)	(-0.073 - 0.021)	(-0.194 - -0.017)
date_c	-0.000	-0.005***	-0.007**	-0.004	-0.024***	0.028
	(-0.001 - 0.001)	(-0.009 - -0.002)	(-0.013 - -0.002)	(-0.014 - 0.005)	(-0.041 - -0.007)	(-0.012 - 0.068)
1.dummy_or#c.date_c	-0.001	0.004	0.002	0.000	0.017*	-0.061***
	(-0.008 - 0.006)	(-0.004 - 0.012)	(-0.004 - 0.008)	(-0.009 - 0.009)	(-0.001 - 0.035)	(-0.099 - -0.023)
Constant	4.054***	3.959***	3.937***	3.944***	3.885***	3.982***
	(3.905 - 4.202)	(3.808 - 4.110)	(3.778 - 4.096)	(3.790 - 4.098)	(3.717 - 4.054)	(3.830 - 4.135)
Observations	1,732	1,202	828	440	228	117
R-squared	0.154	0.144	0.129	0.046	0.048	0.023

Table A.4: National-level Polynomial RDiT estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Model1	Model2	Model3	Model4	Model5	Model6	Model7	Model8
VARIABLES	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag
dummy_or = 1	-0.213***	-0.083***	-0.109***	-0.074*	-0.181**	-0.048**	-0.047	-0.104***
	(-0.312 - -0.114)	(-0.119 - -0.047)	(-0.170 - -0.048)	(-0.155 - 0.007)	(-0.323 - -0.040)	(-0.088 - -0.007)	(-0.106 - 0.012)	(-0.178 - -0.031)
date_c	0.002**	-0.004***	-0.006**	-0.008	0.012	-0.021**	-0.010	0.024
	(0.000 - 0.003)	(-0.006 - -0.003)	(-0.011 - -0.001)	(-0.019 - 0.002)	(-0.004 - 0.028)	(-0.038 - -0.005)	(-0.030 - 0.011)	(-0.030 - 0.077)
1.dummy_or#c.date_c	-0.005***	-0.003	0.016*	0.002	0.024	-0.035	-0.142**	-0.175**
	(-0.009 - -0.002)	(-0.012 - 0.005)	(-0.001 - 0.033)	(-0.050 - 0.054)	(-0.035 - 0.084)	(-0.111 - 0.042)	(-0.260 - -0.023)	(-0.339 - -0.011)
date_csq		-0.000***	-0.000	-0.000	0.001**	-0.002**	-0.000	0.005
		(-0.000 - -0.000)	(-0.000 - 0.000)	(-0.001 - 0.000)	(0.000 - 0.002)	(-0.003 - -0.000)	(-0.003 - 0.002)	(-0.003 - 0.012)
date_cub			-0.000	-0.000	0.000**	-0.000***	-0.000	0.000
			(-0.000 - 0.000)	(-0.000 - 0.000)	(0.000 - 0.000)	(-0.000 - -0.000)	(-0.000 - 0.000)	(-0.000 - 0.001)
date_4				-0.000	0.000**	-0.000***	0.000	0.000
				(-0.000 - 0.000)	(0.000 - 0.000)	(-0.000 - -0.000)	(-0.000 - 0.000)	(-0.000 - 0.000)
date_5					0.000**	-0.000***	0.000	0.000
					(0.000 - 0.000)	(-0.000 - -0.000)	(-0.000 - 0.000)	(-0.000 - 0.000)
date_6						-0.000***	0.000	0.000*
						(-0.000 - -0.000)	(-0.000 - 0.000)	(-0.000 - 0.000)
date_7							0.000	0.000*
							(-0.000 - 0.000)	(-0.000 - 0.000)
date_8								0.000
								(-0.000 - 0.000)
1.dummy_or#c.date_csq		0.000	-0.001*	0.002	-0.011	0.028*	0.079***	0.074***
		(-0.000 - 0.001)	(-0.003 - 0.000)	(-0.006 - 0.009)	(-0.028 - 0.005)	(-0.005 - 0.062)	(0.024 - 0.134)	(0.024 - 0.124)
1.dummy_or#c.date_cub			0.000*	-0.000	0.001	-0.004*	-0.015***	-0.016***
			(-0.000 - 0.000)	(-0.001 - 0.000)	(-0.001 - 0.003)	(-0.009 - 0.001)	(-0.025 - -0.005)	(-0.026 - -0.005)
1.dummy_or#c.date_4				0.000	-0.000	0.000*	0.001***	0.001***
				(-0.000 - 0.000)	(-0.000 - 0.000)	(-0.000 - 0.001)	(0.000 - 0.002)	(0.000 - 0.002)
1.dummy_or#c.date_5					0.000	-0.000*	-0.000***	-0.000***
					(-0.000 - 0.000)	(-0.000 - 0.000)	(-0.000 - -0.000)	(-0.000 - -0.000)
1.dummy_or#c.date_6						0.000**	0.000**	0.000**
						(0.000 - 0.000)	(0.000 - 0.000)	(0.000 - 0.000)
1.dummy_or#c.date_7							-0.000**	-0.000**
							(-0.000 - -0.000)	(-0.000 - -0.000)
1o.dummy_or#co.date_8								0.000
								(0.000 - 0.000)
Constant	4.139***	4.027***	4.015***	4.000***	4.077***	3.984***	4.010***	4.068***
	(4.049 - 4.229)	(4.001 - 4.053)	(3.977 - 4.054)	(3.942 - 4.058)	(4.016 - 4.139)	(3.932 - 4.036)	(3.969 - 4.051)	(3.975 - 4.160)
Observations	135	135	135	135	135	135	135	135
R-squared	0.609	0.769	0.776	0.780	0.811	0.851	0.859	0.866

Table A.5: Regional-level Polynomial RDiT estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Model1	Model2	Model3	Model4	Model5	Model6	Model7	Model8
VARIABLES	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag	ma12_birthlag
dummy_or = 1	-0.232***	-0.077***	-0.024	-0.092***	-0.037*	-0.071**	-0.097***	-0.064**
	(-0.338 - -0.126)	(-0.132 - -0.023)	(-0.080 - 0.032)	(-0.132 - -0.051)	(-0.078 - 0.004)	(-0.137 - -0.005)	(-0.141 - -0.053)	(-0.122 - -0.007)
date_c	-0.000	-0.008***	-0.014***	-0.008	-0.021**	-0.011	0.012	-0.011
	(-0.001 - 0.001)	(-0.011 - -0.005)	(-0.023 - -0.005)	(-0.020 - 0.004)	(-0.036 - -0.005)	(-0.041 - 0.019)	(-0.012 - 0.035)	(-0.051 - 0.029)
1.dummy_or#c.date_c	-0.001	-0.002	0.002	0.019	0.018	0.024	-0.031	-0.009
	(-0.008 - 0.006)	(-0.010 - 0.006)	(-0.009 - 0.013)	(-0.008 - 0.046)	(-0.010 - 0.045)	(-0.005 - 0.054)	(-0.084 - 0.022)	(-0.072 - 0.054)
date_csq		-0.000***	-0.000**	0.000	-0.001**	0.000	0.003***	-0.001
		(-0.000 - -0.000)	(-0.000 - -0.000)	(-0.000 - 0.000)	(-0.002 - -0.000)	(-0.003 - 0.003)	(0.001 - 0.006)	(-0.008 - 0.007)
date_cub			-0.000**	0.000	-0.000**	0.000	0.000***	-0.000
			(-0.000 - -0.000)	(-0.000 - 0.000)	(-0.000 - -0.000)	(-0.000 - 0.000)	(0.000 - 0.000)	(-0.001 - 0.001)
date_4				0.000**	-0.000***	0.000	0.000***	-0.000
				(0.000 - 0.000)	(-0.000 - -0.000)	(-0.000 - 0.000)	(0.000 - 0.000)	(-0.000 - 0.000)
date_5					-0.000***	0.000	0.000**	-0.000
					(-0.000 - -0.000)	(-0.000 - 0.000)	(0.000 - 0.000)	(-0.000 - 0.000)
date_6						0.000	0.000**	-0.000
						(-0.000 - 0.000)	(0.000 - 0.000)	(-0.000 - 0.000)
date_7							0.000**	-0.000
							(0.000 - 0.000)	(-0.000 - 0.000)
date_8								-0.000
								(-0.000 - 0.000)
1.dummy_or#c.date_csq		0.000*	0.001	-0.003*	0.001	-0.005	0.006	0.010
		(-0.000 - 0.001)	(-0.000 - 0.002)	(-0.005 - 0.000)	(-0.004 - 0.006)	(-0.014 - 0.004)	(-0.014 - 0.025)	(-0.011 - 0.030)
1.dummy_or#c.date_cub			-0.000	0.000*	-0.000	0.000	-0.002	-0.002
			(-0.000 - 0.000)	(-0.000 - 0.000)	(-0.001 - 0.000)	(-0.001 - 0.001)	(-0.005 - 0.001)	(-0.005 - 0.002)
1.dummy_or#c.date_4				-0.000*	0.000	-0.000	0.000	0.000
				(-0.000 - 0.000)	(-0.000 - 0.000)	(-0.000 - 0.000)	(-0.000 - 0.000)	(-0.000 - 0.000)
1.dummy_or#c.date_5					-0.000	0.000	-0.000	-0.000
					(-0.000 - 0.000)	(-0.000 - 0.000)	(-0.000 - 0.000)	(-0.000 - 0.000)
1.dummy_or#c.date_6						-0.000	0.000	0.000
						(-0.000 - 0.000)	(-0.000 - 0.000)	(-0.000 - 0.000)
1.dummy_or#c.date_7							-0.000	-0.000
							(-0.000 - 0.000)	(-0.000 - 0.000)
1o.dummy_or#co.date_8								0.000
								(0.000 - 0.000)
Constant	4.054***	3.950***	3.904***	3.933***	3.893***	3.916***	3.957***	3.924***
	(3.905 - 4.202)	(3.801 - 4.098)	(3.745 - 4.063)	(3.766 - 4.101)	(3.726 - 4.061)	(3.761 - 4.071)	(3.793 - 4.120)	(3.766 - 4.081)
Observations	1,732	1,732	1,732	1,732	1,732	1,732	1,732	1,732
R-squared	0.154	0.174	0.177	0.179	0.180	0.180	0.181	0.181