Avendano, Mauricio, Berkman, Lisa F., Brugiavini, Agar and Pasini, Giacomo

The long-run effect of maternity leave benefits on mental health: evidence from European countries

Article (Accepted version) (Refereed)

Original citation:
Avendano, Mauricio, Berkman, Lisa F., Brugiavini, Agar and Pasini, Giacomo (2015) The long-run effect of maternity leave benefits on mental health: evidence from European countries. Social Science & Medicine, 132 . pp. 45-53. ISSN 0277-9536
DOI: 10.1016/j.socscimed.2015.02.037

The long-run effect of maternity leave benefits on mental health: Evidence from European countries

Mauricio Avendano, Lisa F. Berkman, Agar Brugiavini, Giacomo Pasini

PII: S0277-9536(15)00129-X
DOI: doi:10.1016/j.socscimed.2015.02.037
Reference: SSM 9965

Published in: *Social Science & Medicine*

Cite this article as: Avendano M, Berkman LF, Brugiavini A, Pasini G, The long-run effect of maternity leave benefits on mental health: Evidence from European countries, *Social Science & Medicine*, doi:10.1016/j.socscimed.2015.02.037

The long-run effect of maternity leave benefits on mental health: Evidence from European countries

Mauricio Avendano[1,2,3], Lisa F. Berkman[2,3], Agar Brugiavini[4], Giacomo Pasini[4,5]

Affiliations:

1. London School of Economics and Political Science, Department of Social Policy, LSE Health and Social Care, Houghton Street, London, United Kingdom
2. Harvard School of Public Health, Department of Social and Behavioral Sciences, 677 Huntington Avenue, Boston, MA 02115, United States
3. Center for Population and Development Studies, Harvard School of Public Health, 9 Bow Street, Cambridge, USA
4. Ca' Foscari University of Venice, Department of Economics, Venezia, Italy
5. Network for Studies on Pensions, Aging and Retirement, Tilburg, The Netherlands

Corresponding Author:

Mauricio Avendano

Cowdray House, London School of Economics and Political Science, Houghton Street London WC2A 2AE

Tel.: +44 20 7955 7203

E-mail: M.Avendano-Pabon@lse.ac.uk

Word count: 7,764

ACKNOWLEDGEMENTS

We wish to thank Rob Alessie, Danilo Cavapozzi, Mike Hurd, Pilar Garcia Gomez, Jonathan Skinner and Axel Börsh-Supan for valuable comments. Funding for this work came from the National Institute on Aging (grants R01AG040248 and R01AG037398), a Starting Researcher grant from the European Research Council (ERC grant No 263684), and the McArthur Foundation Research Network on Ageing.

This paper uses data from SHARE wave 1 and 2 release 2.6.0, as of November 29 2013 (DOI: 10.6103/SHARE.w1.260 and 10.6103/SHARE.w2.260) and SHARELIFE release 1, as of November 24th 2010 (DOI: 10.6103/SHARE.w3.100). The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5- CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, N° 211909, SHARE-LEAP, N° 227822 and SHARE M4, N° 261982). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions).

ABSTRACT

This paper examines whether maternity leave policies have an effect on women´s mental health in older age. We link data for women aged 50 years and above from countries in the Survey of Health, Ageing and Retirement in Europe (SHARE) to data on maternity leave legislation from 1960 onwards. We use a difference-in-differences approach that exploits changes over time within countries in the duration and compensation of maternity leave benefits, linked to the year women were giving birth to their first child at age 16 to 25. We compare late-life depressive symptom scores (measured with a 12-item version of the Euro-D scale) of mothers who were in employment in the period around the birth of their first child to depression scores of mothers who were not in employment in the period surrounding the birth of a first child, and therefore did not benefit directly from maternity leave benefits. Our findings suggest that a more generous maternity leave during the birth of a first child is associated with a reduced score of 0.38 points in the Euro-D depressive symptom scale in old age.

Keywords: Europe; Maternity leave; depression; social policy; ageing; mental health; international

INTRODUCTION

Social policies can have unanticipated health consequences. Studies on the earned income tax credit, the US welfare reform and the food stamp programme show that although these policies were not motivated by health concerns, they have both negative and positive health externalities (Almond et al., 2011; Bitler et al., 2005; Schmeiser, 2009; Snyder & Evans, 2006). During the second half of the 20th Century, most high-income countries enacted comprehensive maternity leave legislation that provides women the right to a period of job-protected leave around childbirth (Ruhm & Teague, 1998). An extensive literature has examined impacts of these policies on labor market (Dahl et al., 2013; Rossin-Slater et al., 2013; Ruhm, 1998; Ruhm, 2011) and child outcomes (Baker & Milligan, 2008; Berger et al., 2005; Rossin, 2011; Ruhm, 2000, 2011; Staehelin et al., 2007; Tanaka, 2005). However, few studies have examined the impact of maternity leave policies on women's health, with existing studies focusing on health around childbirth (Dagher et al., 2013; Ruhm, 2011; Staehelin et al., 2007).

Maternity leave policies may have long-term effects on mother's health by preventing or reducing the stress around childbirth. New mothers are at increased risk for a range of psychiatric disorders including depression, posttraumatic stress disorder, anxiety and postpartum psychosis(Brockington, 2004). 10% to 15% of mothers experience depression in the postpartum period (Hasin et al., 2005, Wisner et al., 2002, 2009), which may increase vulnerability to subsequent episodes of major depression and other psychiatric disorders in older age (Hammen, 2003; Kessler, 1997). Late life depression is a growing public concern: The Global Burden of Disease report ranks major depressive disorders as the second leading cause of disability(Ferrari et al., 2013). In the United States alone, depression costs $83.1 billion in economic costs (Greenberg et al., 2003). The prevalence of late life depressive symptoms among European women ranges from 18% to 37% (Castro-Costa et al., 2007). Depression leads to impairments in social functioning, quality of life, and increased risk of health problems (McCall & Kintziger, 2013).

Our study examines whether maternity leave benefits lead to enduring benefits in long-term mental health. Initially motivated by concerns for the health of mothers and children, maternity leave policies were first introduced as a prohibition to employers to employ women during pregnancy, but provided

no income or job protection (Ruhm & Teague, 1998). Since the 1960's, maternity leave policies evolved from prohibitions to a time-off work to care for children, combined with job protection for parents. The long-term impact of these policies on women's well-being, however, is poorly understood. We use data from the Survey of Health, Ageing and Retirement in Europe (SHARE) linked to the 'Comparative Maternity, Parental and Childcare Leave and Benefits Database' (Gauthier, 2011). Our identification strategy exploits variation over time across European countries in the enactment of legislation on maternity leave benefits (Gornick & Meyers, 2003). Based on a difference-in-difference approach, results provide evidence of the impact of paid maternity leave around the birth of a first child on late-life depression.

BACKGROUND

Maternity leave and maternal outcomes

A growing literature examines the impact of maternity leave on labor market outcomes, documenting effects on job continuity, wage level and growth, labor market attachment and employability (Brugiavini A et al., 2012; Dahl et al., 2013; Klerman JA & Leibowitz A, 2000; Klerman & Leibowitz, 1999; Rossin-Slater et al., 2013). Studies have also examined impacts of maternity leave policies on mother's health around childbirth. In a systematic review (Staehelin et al., 2007), four of six studies reported positive associations between length of maternity leave and post-partum mental health. Recent studies use variation in policies to study the health effects of maternity leave. Exploiting cross-sectional variation in policies across US states, Chatterji & Markowitz find that longer leave is associated with reductions in depressive symptoms (Chatterji & Markowitz, 2005; Chatterji & Markowitz, 2012). Dagher & McGovern exploit variation in employer policies and find that increases in leave duration are associated with decreased depressive symptoms until six months after childbirth (Dagher et al., 2013). On the other hand, Baker & Milligan find no effect of extensions of paid maternity leave in Canada on maternal or child health (Baker & Milligan, 2008). Two of these studies focus on the United States, where rights to maternity leave are short and unpaid, and all three studies examine relatively short-term effects. The large expansion in paid maternity leave

benefits in Europe during previous decades offers a unique source of variation to explore the long-run impact of maternity leave on mothers' mental health.

Maternity Leave and late life mental health

Two theories from psychology provide the basis to link women's mental health to their employment and fertility decisions(Marshall & Barnett, 1993). The 'scarcity hypothesis' (Coser L, 1974; Gooede, 1960; Slater, 1963) postulates that the competing demands from work and family lead to role overload, which may give rise to additional stressors and generated a process of 'stress proliferation'(Frone et al., 1997; Mullen et al., 2008; Pearlin et al., 2005). Maternity leave may relieve the stress from role overload during childbirth, but leave entitlements may also incentivize mothers to maintain multiple roles, thus increasing the potential for stress. Alternatively, the 'expansion' or 'enhancement' hypothesis (Marks, 1977; Sieber, 1974) posits that multiple roles enhance well-being by increasing sources of identify, self-esteem and resources to cope with multiple demands. These benefits from work may offset the stress associated with combining family and work roles(Grzywacz & Bass, 2003). Increasing research supports the notion that participation in multiple work and family roles has positive effects on mental health(Mullen et al., 2008).

Maternity leave policies may also have indirect effects on mother's mental health. A period of leave shortly after birth may improve mother-child relationships and reduce the risk of later disorders in children (Brockington, 2004), which may in turn improve maternal well-being in older age. Women with a prior episode of depression are more likely to experience divorce and marital difficulties, and to have a spouse with psychiatric disorders (Hammen, 2003). Maternity leave benefits may also influence employment and lifetime earnings, which may generate positive externalities on late-life mental health.

DATA AND METHODS

SHARE is a cross-national panel survey designed to provide comparable information on the health, employment and social conditions of a representative sample of the European population aged 50+. Samples were drawn in Northern Europe (Sweden and Denmark), Western Europe (Austria, France,

Germany, Switzerland, Belgium, and the Netherlands), Southern Europe (Spain, Italy and Greece) and Eastern/Central Europe (Poland and Czech Republic)(Borsch-Supan et al., 2013). Our analysis focuses on Western European countries. We excluded Poland, Czech Republic and women living in East Germany before 1989: women in these countries were exposed to a system of full, but not freely chosen, employment (Gal & Kligman, 2000), so that maternity leave decisions were not comparable to those of women in Western countries. Sample size, response rates and attrition rates are summarized in Appendix Table A1. Response rates for the first interview in 2004/5 were 62% on average, although there were differences between countries. Individual retention rates for wave 2 in 2006/7 were 73%, while retention rates were 77% for wave 3 in 2008/9 (Borsch-Supan et al., 2013; Börsch-Supan & Jürges, 2005; Schröder, 2011).

Our measure of depressive symptoms is based on the EURO-Depression (Euro-D) scale, a standardized measured designed for international comparisons of depressive symptoms in Europe. Participants are asked whether during the past month they have experienced any of a list of 12 symptoms: depression, pessimism, death wish, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment and tearfulness (Prince et al., 1999). The score ranges from zero to 12, with higher scores indicating higher levels of depressive symptoms. A score higher than three is suggested as a predictor of depression caseness(Castro-Costa et al., 2007; Prince et al., 1999). Participants were assigned as outcome the Euro-D score in the first wave they were interviewed in SHARE (either 2004/5 or 2005/6).

Data on maternity episodes came from the 2008/09 life history retrospective assessments (Brugiavini A et al., 2013; Schröder, 2011). Using the life-grid History Event Calendar, SHARE participants were asked to report each paid job that lasted for 6 months or more since leaving full-time education (or since the first job for those without any schooling). Participants could report up to 20 job episodes, for each of which they reported several details including the year the job started and ended; occupation and industry; whether job was part- or full-time; and the reasons and duration of any gaps between jobs. As part of the fertility life-history assessment, participants were also asked to report details of each natural child including date of birth, gender and year of death if child had deceased.

Subsequently, participants were asked whether and for how long they had stopped working when each child was born. We then derived a panel of maternity leave and job episodes for each respondent. As expected, labor market participation around childbirth was heterogeneous across countries, ranging from around 60% to less than 40% in countries such as Italy and Spain (Figure 1).

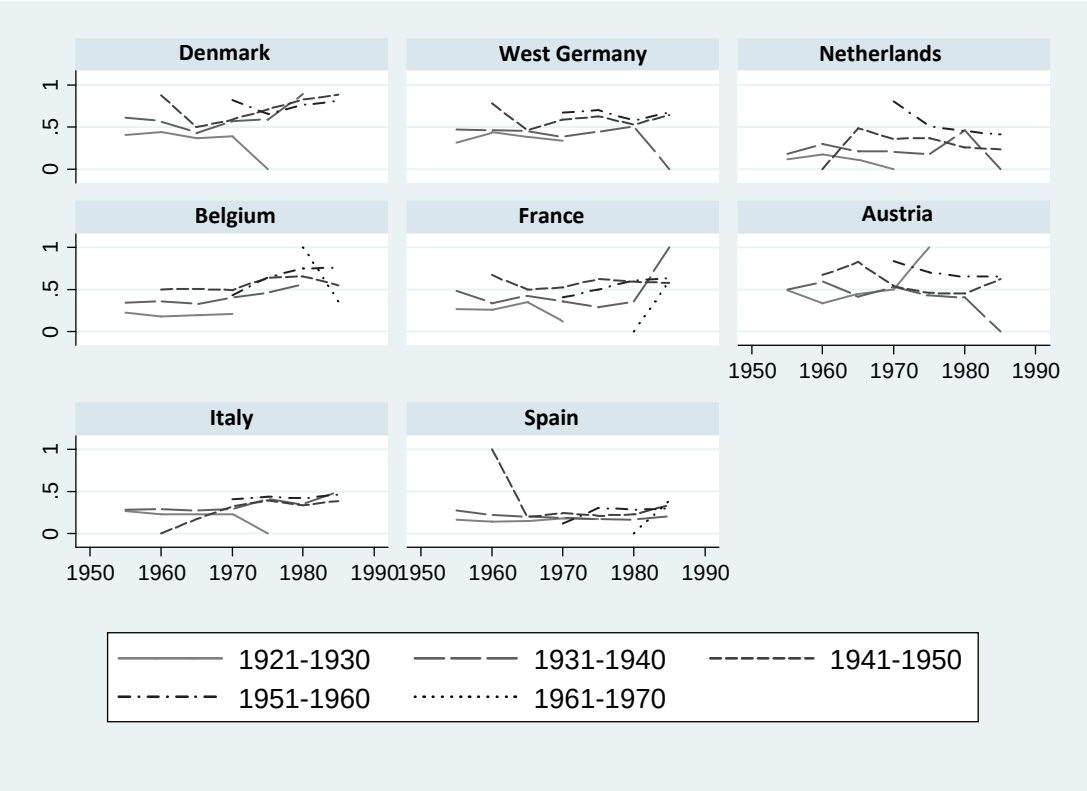

Figure 1

In addition to job and fertility histories, SHARE also included measures of educational attainment, marital status, income, limitations with activities of daily living (ADL) and instrumental activities of daily living (IADL), the Global Assessment of Limitations Index (GALI) item, years of smoking, drinking behavior and number of miscarriages. Detailed variable definitions are provided in Table A2, while Table 1 provides descriptive statistics of all variables.

Table 1: Descriptive statistics

	mean	Sd	min	max
Euro-D depression scale	2.827	2.404	0	12
>3 Euro-D symptoms	0.451	0.498	0	1
Age at interview	66.543	10.528	36	100
Elementary education	0.350	0.477	0	1
High school education	0.355	0.478	0	1
College education	0.187	0.390	0	1
Married/Cohabiting	0.952	0.214	0	1
Number of children	2.533	1.392	1	14
Age at childbirth	24.774	4.493	12	47
Total tenure in the labor market	22.050	15.503	0	60
Log of permanent income	8.505	0.719	6	10
Activities of daily living (ADL)	0.217	0.798	0	6
Instrumental ADL's (IADL)	0.384	1.074	0	7
Limitations with activities	0.433	0.496	0	1
Years smoking	7.731	14.053	0	70
Days drinking in a month	7.894	11.046	0	30
Number of miscarriages	0.074	0.345	0	5
Contraceptive pill available*	0.381	0.486	0	1
Abortion legal*	0.176	0.380	0	1

Maternity leave policy data

We supplement SHARE with key characteristics of legislation on maternity leave in each country over a 50-year span. Data came from the Comparative Maternity, Parental, and Childcare Leave and Benefits Database (Gauthier, 2011). We follow the approach of earlier studies (Ruhm, 1998; Ruhm, 2000) and define maternity leave as the period granted to mothers in connection with childbirth, which includes the period of leave immediately prior and after childbirth. This definition is restrictive: it excludes parental and childcare leave not directly linked to childbirth. However, it allows us to focus on a well-defined component of maternity leave policy that is roughly comparable across countries. Features of extended parental leave are very diverse across countries and there is no consensus on how to operationalize these measures. Moreover, earlier studies have focused primarily on the impact of parental leave around childbirth on mothers, while the evidence for any effects of extended leave after the childbirth period ends remains controversial.

We exclude Sweden from our analysis, because legislation did not distinguish leave around childbirth from extended parental leave. We also exclude Switzerland, Greece and the Netherlands because maternity leave benefits in these countries did not vary during the years in which SHARE women had their first child, resulting in no variation for identification. In sensitivity analyses, however, we found that including these countries yielded very similar results as presented here (Appendix Table A5, columns 1 and 2).

Table A3 provides summary statistics of the two policy variables used: duration of leave in weeks and percentage of past wages replaced during maternity leave (replacement rates). Maternity leave length in most countries ranges from 12 to 18 weeks, with the exception of Italy, where it ranges from 17 to 24 weeks. Most countries offer benefits close to 100% of the previous wage, with the exceptions of Belgium and Denmark. In order to combine the two dimensions along which the maternity leave policy varies, we follow Ruhm (2000) and multiply number of weeks by replacement rates to obtain a summary indicator of generosity defined as the number of weeks of full wage leave (FWW) provided to mothers.

Figure 2 shows the value of FWW from 1960 to 1994, the period during which women in SHARE reported the birth of a child. The figure reveals a switch from a less generous, "low FWW regime", to a more comprehensive, "high FWW regime" over time. While FWW does not allow us to distinguish independent effects of duration and monetary benefit generosity, combining duration and monetary benefit into a single indicator allows us to narrow down the policy to a uni-dimensional indicator that effectively identifies these shifts in regime generosity within each country.

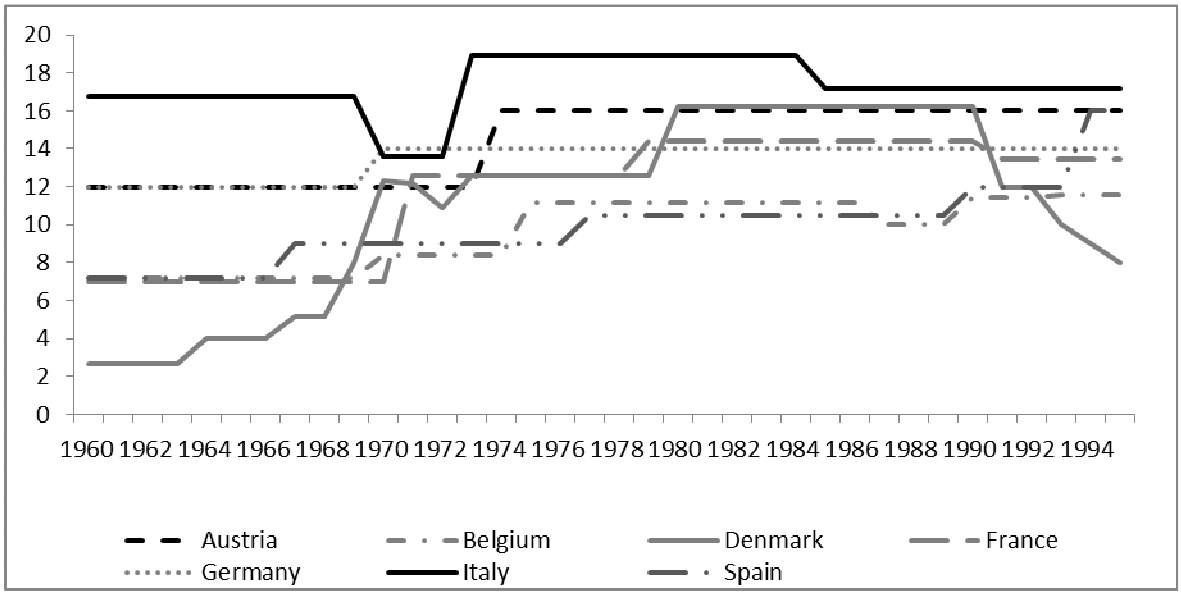

Figure 2

Empirical strategy

Between the 1960s and mid-1990's, policies other than maternity leave such as unemployment insurance and pensions also increased in coverage and generosity. Naïve comparisons of mental health among women exposed to different policy regimes may thus simply capture welfare generosity at large or cohort effects, rather than the specific impact of maternity leave policies on women's mental health. To isolate the impact of maternity leave legislation, therefore, we use a difference-in-differences (DiD) approach.

The rationale for our DiD approach is as follows: we compare outcomes of women who were in employment in the period around the birth of their first child (the treated group) to outcomes of women who were not in employment in the period around the birth of their first child (the control

group). The latter group of women was not eligible for maternity leave benefits, and therefore serves as control group. Table A6 shows that, as expected, women in these two groups differ along several characteristics, which precludes any direct comparison of depressive symptoms between them. A difference in differences approach aims to control for these underlying differences between treatment and control by comparing trends –rather than levels- in depressive symptoms between treatment and control. The DiD estimate is thus the difference between women exposed to comprehensive vs. less generous maternity leave benefits at the time of first childbirth, net of differences in depression scores between treated and control women. Our assumption is that this double difference corresponds to the impact of maternity leave policies on depression scores, because it captures the change in the difference between treated and control as a result of the policy.

The estimation is formalized by equation (1):

$$eurod_{ic} = \beta_0 + \beta_1 work_{ict} + \beta_2 fww_{ct} + \beta_3 work_{ict} * fww_{ct} + \boldsymbol{x}' \beta_X + \varepsilon_{ict} \quad (1)$$

Where $eurod_{ic}$, is the Euro-D score; *work* represents a variable that takes 1 for women in the treatment group (employed in the period around childbirth) and 0 for women in the control group (not employed in the period around childbirth); *fww* takes 1 if maternity leave coverage in country *c* in year of childbirth *t* was comprehensive and 0 otherwise; work*fww is an interaction term between employment status at childbirth and FWW; and X represents a vector of control variables. The coefficient of the interaction term β_3 is exactly the double difference computed at the mean value of the outcome.

A key assumption in the DiD approach is that women do not self-select themselves into treatment or control as a result of the policy. This implies that women do not change their employment status on the basis of the maternity leave policy at childbirth. In order to reduce the impact of self-selection, our primary analysis focuses on births that took place between ages 16 to 25 years, based on the rationale that at young ages women are less likely to self-select into treatment and control. At young age, women face relatively small losses in life-time income as a consequence of maternity interruptions regardless of the policy in place, because wages at young ages are lower than at later ages, and

women at young age have a longer working life ahead to recover from wage losses due to job interruptions.

We did not restrict the sample to mothers who had only one child as this was a small and selective group (415 women, less than 9% of the analytical sample). A concern is that mothers who had more than one child may have been exposed to different regimes in later births. Nevertheless, this would only bias our results if later maternity leave policies affected outcomes in a different way for the treatment and control groups. Finally, we also perform robustness checks that restrict analyses to women who did not change their labor market status two years before childbirth, thus minimizing the impact of self-selection.

Because our model is a country fixed effect, we need variability within countries in generosity of benefits. We therefore defined coverage of maternity leave to be limited or comprehensive on the basis of observed changes in FWW within each country. Table A4 reports the number of first maternity episodes occurring between age 16 and 25 by country of residence against FWW in place at childbirth. The policy variable is defined as a dummy that takes the value of 1 if FWW is larger than 12 weeks in Denmark, Germany, France and Austria; larger than 16 weeks in Italy; and larger than 8 weeks in Belgium and Spain. Country specific thresholds imply that women who gave birth to their first child in different countries and exposed to the same FWW indicator can be considered as being exposed to different maternity leave regimes. Policy changes within each country have a discrete nature: FWW must be interpreted as an indicator that distinguishes less generous from more comprehensive maternity leave benefits, rather than a continuous policy variable. In order to check that results do not depend on the specific categorization into low and high FWW regime, we estimate the model pooling all countries, but also on subsets of countries for which the threshold is the same: Denmark, Germany, France and Austria on the one hand, and Belgium and Spain on the other. The limited sample size does not allow estimation separately for each country.

Regressions include also a set of basic determinants of mental health: a quadratic in age; educational level; marital status; number of children throughout life; and cohort dummies. Moreover, we include a

set of physical health measures (ADL, IADL, and GALI disability) and two measures of health behavior (years spent smoking and alcoholic drinks per month). Given common trends towards more generous welfare states, even controlling for cohorts, our policy variable may simply proxy trends in social spending in each country. Therefore, we add a full set of country-specific linear trends in year of childbirth.

Ethical Approval

The SHARE survey received full ethical approval from the internal review board (IRB) at the University of Mannheim (Germany).

RESULTS

Figure 3 shows the association between the policy indicator, FWW, and depression measured on the 12-point Euro-D scale, separately for women in employment and women not in employment in the period surrounding childbirth. While there is not a clear positive or negative relationship, the difference between the two groups at given values of the policy indicator widens as maternity leave become more comprehensive. In order to better appreciate this, Figure 4 directly plots the difference between mothers in the treated and control groups according to FWW: at low FWW levels, women not in employment in the period surrounding childbirth exhibit lower depression scores than women in employment in the period around childbirth, but this difference is reversed as FWW increases above 10.

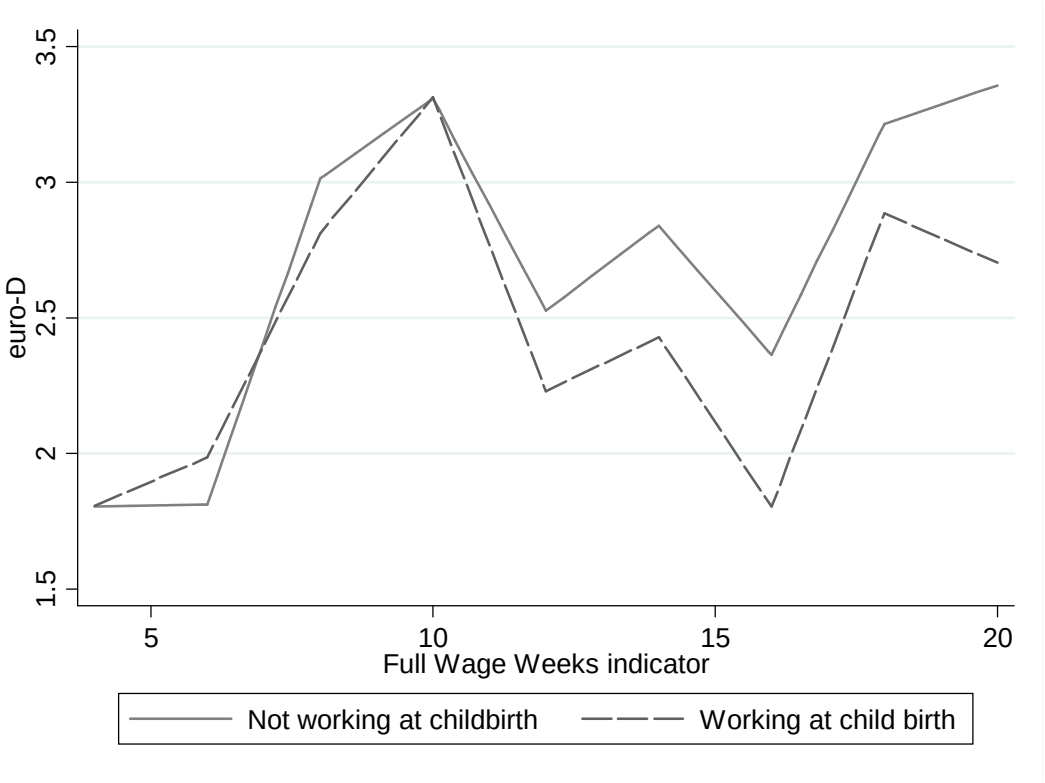
euro-D
1.5
2
2.5
3
3.5
5
10
15
20
Full Wage Weeks indicator
Not working at childbirth
Working at child birth

Figure 3

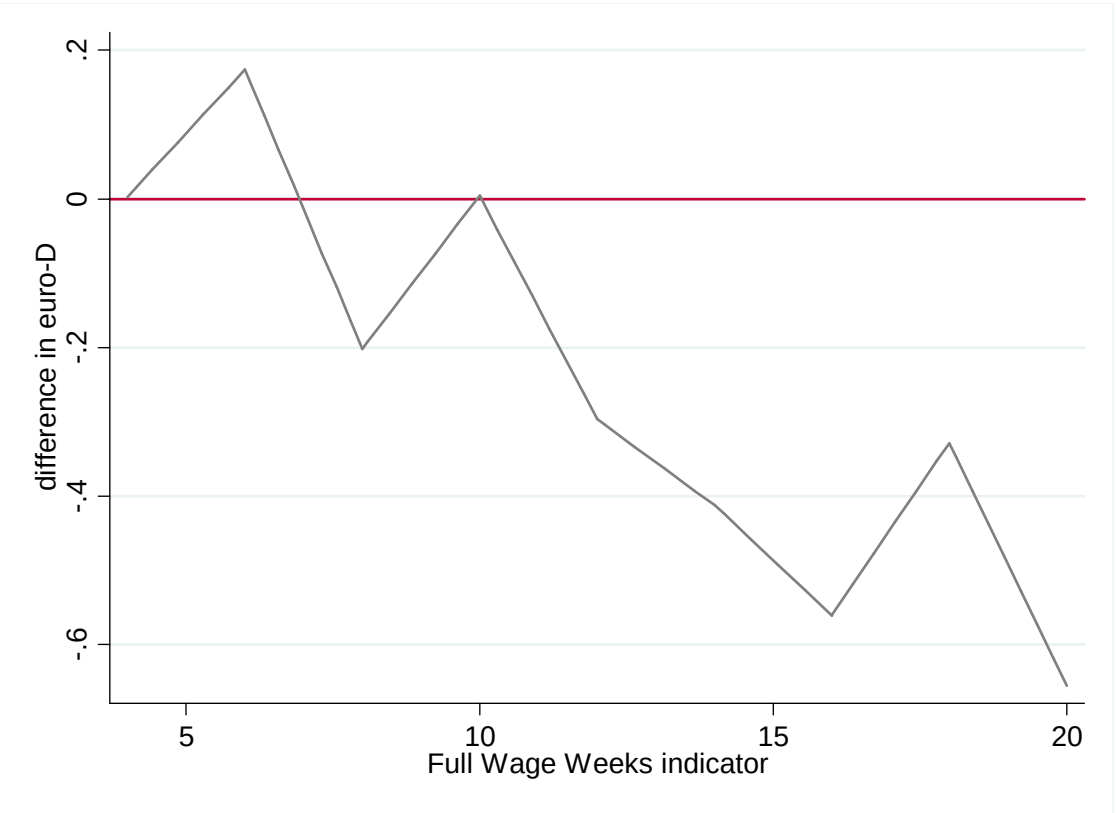
difference in euro-D
-.6
-.4
-.2
0
.2
5
10
15
20
Full Wage Weeks indicator

Figure 4

Table 2 contains the estimation results. The first column reports estimates with all countries pooled; the second restricts the sample to Spain and Belgium, countries in which the threshold between low and high FWW regime is set at eight full wage weeks of maternity. The third column reports estimates based on data for Germany, Austria, Denmark and France, where the threshold to distinguish limited from comprehensive benefits is set at 12 months.

Table 2: Difference-in differences estimation: Euro-Depression scores and full wage weeks of maternity leave

	(1)	(2)	(3)
In employment during period	0,178	0,192	0,165*
surrounding childbirth	(0,118)	(0,213)	(0,097)
comprehensive maternity leave	0,307	0,177	0,530*
	(0,190)	(0,342)	(0,273)
In employment around childbirth*comprehensive	-0,385**	-0,653**	-0,468**
maternity leave	(0,171)	(0,323)	(0,225)
Age	0,07	0,025	0,235
	(0,190)	(0,338)	(0,165)
Age squared	-0,001	-0,001	-0,002
	(0,002)	(0,003)	(0,001)
High school education	-0,145	-0,033	-0,250**
	(0,104)	(0,174)	(0,107)
College education	-0,186	0,04	-0,323**
	(0,134)	(0,263)	(0,133)
Married/Cohabitation vs. other	-0,413**	0,655	-0,536***
	(0,192)	(0,480)	(0,173)
Number of children	0,066*	0,057	0,045
	(0,037)	(0,063)	(0,039)
1931-1940 cohort	.	.	-0,132
	.	.	(0,205)
1941-1950 cohort	-0,022	0,509	0,019
	(0,203)	(0,409)	(0,287)
1951-1960 cohort	0,31	1,160**	0,028
	(0,293)	(0,568)	(0,509)
1961-1970 cohort	0,2	1,945*	0,946
	(0,621)	(1,036)	(0,781)
1971-1980 cohort	0,324	.	.
	(0,884)	.	.
Activities of Daily Living (ADL) score	0,433***	0,440**	0,238*
	(0,123)	(0,198)	(0,138)
Instrumental Activities of Daily Living	0,535***	0,603***	0,401***
(IADL) score	(0,102)	(0,176)	(0,109)

Limitations with activities (GALI)	1,142***	1,128***	1,077***
	(0,093)	(0,178)	(0,092)
Years smoking	0,004	0,004	0,005*
	(0,003)	(0,005)	(0,003)
Days drinking in a month	0,001	0	-0,002
	(0,004)	(0,007)	(0,004)
Year of first childbirth	-0,094***	-0,119**	-0,051*
	(0,028)	(0,046)	(0,029)
N	2857	908	2624

Notes: Stars represent statistical significance: * p<0.10. ** p<0.05. *** p<0.01. Standard errors reported in parenthesis are robust to heterskedasticity. Excluded cohort is 1931-1940 in columns (1) and (2), 1921-30 in column (3). All estimates include first maternity episodes which took place when the mother was between 16 and 25 years old. Comprehensive maternity leave takes value 1 if the value of FWW (full wage weeks) is larger than 12 weeks in Denmark, Germany, France and Austria; larger than 16 weeks in Italy; and larger than 8 weeks in Belgium and Spain. Column (1) includes all countries. Column (2) includes only Belgium and Spain. Column (3) includes only Denmark, Germany, France and Austria. In addition to variables in table all models include a constant, country fixed effects and country-specific linear trends on age at first birth

Estimates in column (1) and (2) indicate that a more generous maternity leave policy at first childbirth is unrelated to depression scores among mothers not in employment, while estimates in column (3) suggest that more generous benefits are associated with an increase of 0.52 points in Euro-D scores among mothers not in employment. Among mothers in employment around childbirth, the sum of coefficients for being in employment around childbirth and the interaction is not significant in column (1), and points to a reduction of depression scores by 0.46 points in column (2) and by 0.29 points in column (3), both significant at 10%. These associations, however, do not identify a causal effect as they may reflect changes in variables other than maternity leave policies. The net effect of comprehensive maternity leave policies must be evaluated focusing on the interaction term, which corresponds to the difference-in-differences estimate.

Focusing on the first column, the difference-in-differences estimate of the effect of the policy, net of common unobserved differences between treatment and control, is a reduction of the depression Euro-D score of 0.38 points, significant at the 5% level. Impact increases to 0.65 in column (2) and to 0.47 points in column (3) when we restrict the sample to countries with common thresholds separating low and high FWW regimes.

Turning to other variables in the model, indices of physical limitations are strongly associated with higher depression scores. Married and cohabiting women have lower depression scores than single women in column (1) and (3), while the association is not significant in column (2). The total number of children is associated with increases in the depression score only in column (1). Age and cohort dummies do not predict depression scores, but their statistical significance may be reduced due to their correlation with country specific trends in year of childbirth.

Sensitivity analyses

Response and attrition rates were high in some countries and they may be selective on relevant characteristics potentially leading to bias. To account for this, we carried out analyses including country fixed effects, and we repeated the estimation on different sets of countries (columns (2) and (3) of table 2, columns (1) and (2) of table A5). In all cases, results are in line with our baseline specification. Estimates in table 2 consider only mothers who gave birth to their first child before turning 26 years old in order to limit the impact of self-selection into treatment and control. Nevertheless, the sample of women who had their first child when they were younger than 26 may be selective. In column (1) of table 3 we run the same model without any restriction on age of the mother at first birth. The coefficient of interest is almost identical to that for the baseline specification, and it is precisely estimated: moving from a limited coverage maternity leave benefit to a more comprehensive one leads to a reduction of 0.35 points on the Euro-D scale, significant at the 1% level.

Table 3: Robustness checks: Euro-Depression scores and full wage weeks of maternity leave

	(1)	(2)	(3)	(4)	(5)
In employment around childbirth	0,118	0,310**	0,244*	0,158	0,274**
	(0,096)	(0,133)	(0,136)	(0,135)	(0,135)
Comprehensive maternity leave	0,313**	0,395*	0,391*	-0,221	0,258
	(0,133)	(0,212)	(0,212)	(0,411)	(0,167)
In employment around	-0,351***	-0,503***	-0,360*	-0,375	-0,358**
birth*comprehensive maternity	(0,131)	(0,192)	(0,192)	(0,245)	(0,169)
Age	0,189*	0,024	0,045	0,093	0,052
	(0,108)	(0,203)	(0,202)	(0,433)	(0,211)
Age squared	-0,002*	-0,001	-0,001	-0,001	0
	(0,001)	(0,002)	(0,002)	(0,003)	(0,002)
High school education	-0,186**	-0,159	-0,099	-0,229*	-0,108
	(0,082)	(0,114)	(0,112)	(0,138)	(0,103)
College education	-0,390***	-0,152	-0,141	-0,217	-0,179
	(0,096)	(0,148)	(0,142)	(0,176)	(0,133)
Married/Cohabitation or Single	-0,14	-0,458**	-0,429**	.	-0,132
	(0,156)	(0,192)	(0,200)	.	(0,116)
Number of children	0,048	0,036	0,046	0,025	0,027
	(0,030)	(0,042)	(0,040)	(0,050)	(0,039)
1931-1940 cohort	-0,267	-0,119	-0,309	-0,538	-0,127
	(0,353)	(0,933)	(0,928)	(0,394)	(1,332)
1941-1950 cohort	-0,283	-0,182	-0,389	-0,463	-0,008
	(0,428)	(0,958)	(0,951)	(0,324)	(1,388)
1951-1960 cohort	-0,051	0,161	-0,179	.	0,132
	(0,456)	(0,869)	(0,862)	.	(1,315)
1961-1970 cohort	-0,25	-0,272	-0,147	.	-0,046
	(0,552)	(0,629)	(0,594)	.	(0,963)
1971-1980 cohort	0,3	.	.	.	.
	(0,687)	.	.	.	.
Activities of daily living (ADL) score	0,251**	0,429***	0,383***	0,512***	0,21
	(0,102)	(0,133)	(0,130)	(0,152)	(0,142)
Instrumental activities of daily living	0,603***	0,520***	0,520***	0,583***	0,628***
(IADL) score	(0,076)	(0,128)	(0,108)	(0,137)	(0,126)

Limitations with activities (GALI)	1,250***	1,037***	1,214***	1,088***	1,372***
	(0,071)	(0,100)	(0,099)	(0,120)	(0,098)
Years smoking	0,006***	0,002	0,004	0,003	0,002
	(0,002)	(0,003)	(0,003)	(0,004)	(0,003)
Days drinking in a month	0	0,003	0,001	0,004	0
	(0,003)	(0,004)	(0,004)	(0,005)	(0,004)
Year of child birth	-0,023*	-0,106***	- 0,118***	-0,023	.
	(0,012)	(0,030)	(0,030)	(0,050)	.
Years spent working	.	-0,007	.	.	.
	.	(0,004)	.	.	.
Log of permanent income	.	-0,025	.	.	.
	.	(0,072)	.	.	.
N obs	4860	2419	2481	1530	2630

Notes: Stars represent statistical significance: * p<0.1. ** p<0.05. *** p<0.01. Standard errors reported in parenthesis are robust to heterskedasticity. Excluded cohort is 1921-1930. Germany is the excluded country specific trend in year of childbirth. Austria trend as well as married/cohabiting dummy in column (4) are dropped since there are no observations of cohabiting women nor Austrian women in the specific sample. Column (1) includes all first maternities. Comprehensive maternity leave takes value 1 if the value of FWW (full wage weeks) is larger than 12 weeks in Denmark, Germany, France and Austria; larger than 16 weeks in Italy; and larger than 8 weeks in Belgium and Spain. Column (2) includes all women aged 16-25 at childbirth and adds years spent working until 2010 and logarithm of permanent income. Column (3) includes all women aged 16-25 at childbirth who were either employed since at least two years prior to childbirth, or were not employed since two years prior to childbirth. Column (4) includes all women aged 16-25 at childbirth with an unplanned maternities index with a value above the 75% percentile of the distribution in the full sample of mothers. Column (5) uses as control group women working at least 5 years between age 16 and 25, with no maternities within this age range. FWW for the control group refer to FWW relative to age 25 of each woman in the group. In addition to variables in table, all models include a constant, country fixed effects and columns (1) to (4) a country-specific linear trends on age at first birth as well.

Specifications in Table 2 do not include socioeconomic status measures among regressors as they may be directly influenced by maternity leave generosity. On the other hand, controlling for these factors may provide estimates of the impact of the policy net of labor market and late-life income effects. In Table 3, the model in column (2) includes the total number of years worked until 2010 and the (logarithm of) permanent income of the mothers, evaluated at the time of 2008/9 interview and computed according to Brunello et al (Brunello G et al., 2012). The coefficient of interest (the interaction between employment status and FWW) increases to 0.50 and it is not statistically different from the coefficient in the baseline estimation in column (1) of Table 2.

In column (3) of table 3, we restrict the sample to mothers who did not change labor market status two years prior to childbirth. This enables us to further control for self-selection, i.e. to consider only women whose decision to enter or leave the labor market was not directly influenced by a change in the maternity leave policy regime. Again, the coefficient of interest is not statistically different from our baseline estimate. Still, women could plan timing of childbirth very early, and account for generosity of the maternity leave even in this restricted time period. In column (4), we include only maternity episodes likely to be unplanned. Although our data do not include information on whether maternities are planned, we are able to build an index of the likelihood that a given childbirth is unplanned. To do this, we construct a dummy which takes the value of 1 if a given childbirth took place in a country and year where abortion was not legalized; a second variable which takes the value of 1 if the contraceptive pill was not available; a third variable that takes the value of 1 if the mother had miscarriages in her life; and a fourth variable that takes the value of 1 if the mother was younger than 18 at childbirth. Finally, we create a fifth variable that takes the value of 1 if the mother had no partner at childbirth. The unplanned maternity index is the first principal component from this set of indicators. We then run the regression of interest restricting the sample to women in the top quartile of the distribution of the unplanned maternity index. Again the coefficient of interest in column (4) is not statistically different from the one in the baseline regression, but it is imprecisely estimated due to reduced sample size.

A concern is that our control group of women not in employment in the period around childbirth might not offer a good counterfactual for our treatment group, given the potential selection processes associated with labor market inactivity. Column (5) shows estimates using as control group women who were employed at least 5 years between ages 16 to 25 but did not have children in this period. The rationale is that this group of employed women is less selected than the sample of women not employed in the period surrounding childbirth, yet they would not have benefitted from maternity leave benefits. We assigned as treatment to these women the number of weeks of maternity leave they would have been eligible for at age 25 (choosing average of FWW faced between age 20 and 25 leads to similar results). Results suggest that estimates are robust to the use of this alternative control

group: the difference in differences estimate suggests that a more comprehensive maternity leave at first childbirth is associated with significantly less depressive symptoms in older age (-0.358, p <.05). The estimate is very similar to that obtained in our original specification (-0.385, p<.05, Table 2), and the two estimates are not statistically different from each other.

Common trend assumption

A difference-in-differences estimator relies on the common trend assumption: given a policy implemented at a certain date, the increase or decline in the mean outcome among the treated observed after this date in the treatment group would have been equal to the increase or decline observed in the control group, had the policy not been implemented. Such hypothesis cannot be directly tested since the counterfactual is not observable. The usual approach in the literature, if the outcome is observed repeatedly over time, is to check whether the outcome variable follows parallel trends over the period except in the year of the reform. In our specific case, first maternities by definition can be observed only once per individual. Still, depression among working and non-working mothers should follow parallel trends along the year of first childbirth, except for the years in which there is a policy change. Figure 5 reports mean Euro-D scores in the treated and in the control group by year of birth of the first child. The solid lines report the actual depression scores, the dashed lines the predicted scores regressing Euro-D on the full set of controls $\boldsymbol{x}$ in equation (1). Looking at trends both in actual and fitted values, the graphical test does not highlight any clear difference between the treated and control group. The hypothesis can be statistically tested running a regression of Euro-D on a polynomial in the year of childbirth fully interacted with an indicator for employed mother, and then running a Wald test on the joint significance of all the interaction terms. The test accepts the null of equality of the trends.

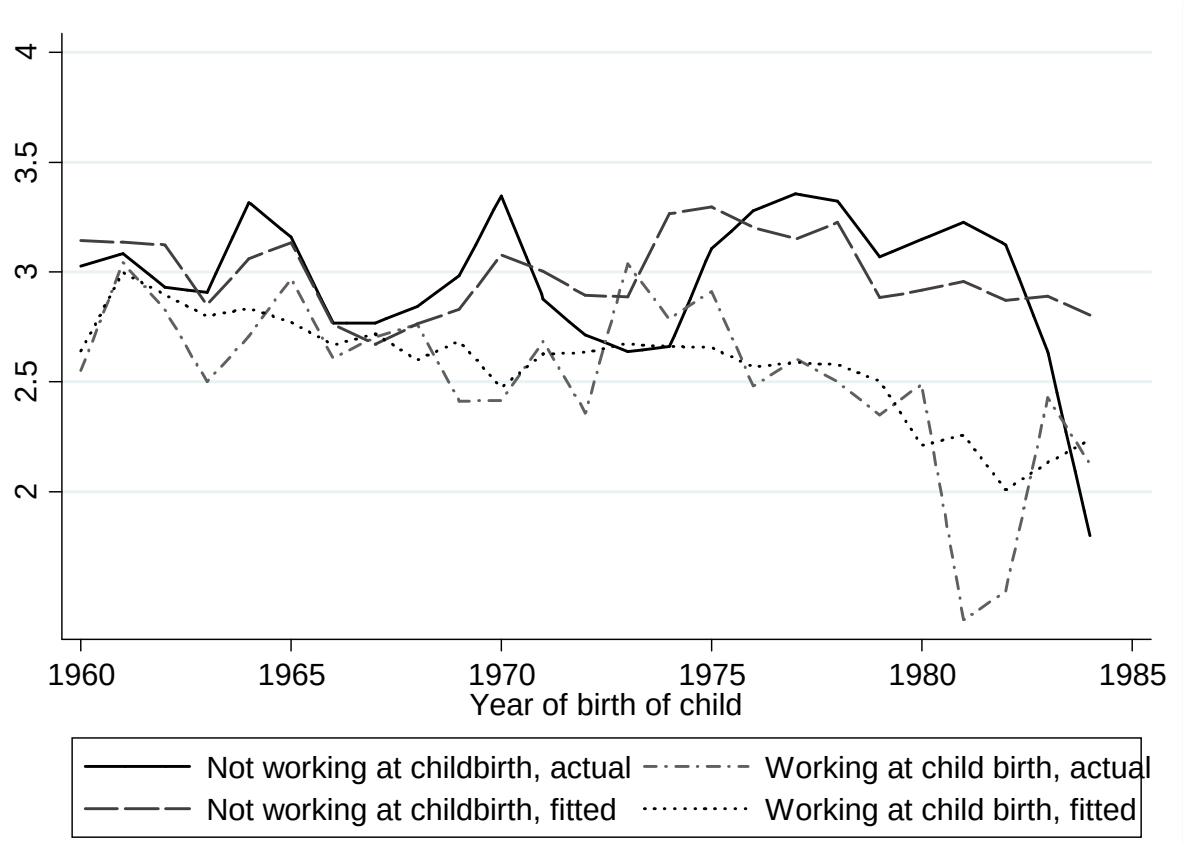

Figure 5

CONCLUSION

Depression is a leading cause of disability in older age(Ferrari et al., 2013). Our results suggest that depression in older age is linked to maternity leave policies during the critical period of the birth of a first child. Our findings suggest that maternity leave benefits, which are designed to provide mothers with a job-protected period around childbirth, have beneficial health effects that extend beyond those documented in earlier studies on labor market careers, wage level and growth, labor market attachment and employability (Brugiavini A et al., 2012; Klerman JA & Leibowitz A, 2000; Klerman & Leibowitz, 1999; Rossin-Slater et al., 2013).

Estimates from earlier studies, obtained with different methodologies, report that the short-term effects of maternity leave on depression are comparable or larger to those we observe. For example, using US data and an instrumental variable approach, Chatterij and Markowitz(Chatterji &

Markowitz, 2012) find that a maternity leave of less than 12 weeks is associated with an increase of 0.79 points in the 12-item CES-D depressive symptom score. Similarly, based on an instrumental variable model, Dagher et al (Dagher et al., 2013) find that mothers who took 12 weeks of leave had a mean score of 5.24 in a 10-item Edinburgh Postnatal Depression Scale, vs. 3.30 for mothers who returned to work within 12 months after child birth (score ranges from 0-30). These studies examine impacts of unpaid maternity leave, while our study examines paid maternity leave.

We found that a comprehensive maternity leave is associated with a reduction of 0.38 points in the Euro-D score. To provide a sense of the magnitude of this effect, we estimated that this would correspond to a Cohen's d of 0.15. This would be considered a small effect according to Cohen's conventional criterion of 0.20. To provide a measure of relative effect size, we also estimated relative risks of depressive symptomatology and found that a comprehensive maternity leave reduces risks of reporting more than three symptoms by 18% (relative risk comparing women with a comprehensive vs. a less comprehensive maternity leave=0.82, 95% Confidence Interval 0.70, 0.96). If this number is estimated with probit or linear probability models rather than a Poisson, the marginal reduction of the probability of reporting three or more symptoms is reduced by 10% (table A5, column 4). Overall, although the magnitude of the effect is relatively small, the effect is comparable to effects of other social and physical health variables, such as being married or having a limitation with activities of daily living (Table 2 and Table A5).

A key question relates to mechanisms that account for the effects of maternity leave on mother's long-run mental health. A possible explanation is that maternity leave benefits reduce the risk of mental health problems shortly after childbirth, which may in turn reduce the risk of future episodes of depression in older age. There is some support for the hypothesis that maternity leave benefits improve mental health outcomes around the period of birth (Chatterji & Markowitz, 2012; Dagher et al., 2013; Ruhm, 2011; Staehelin et al., 2007), potentially influencing mental health in the long-run.

The main policy implication of our paper is that maternity leave legislation in Europe has brought important long-run mental health benefits for mothers, which should be taken into account when

considering the impact of maternity leave. The current financial crisis, for example, has sparked debates on costs and benefits of maternity leave benefits, with some countries such as Czech Republic, the Netherlands, Ireland and Lithuania implementing benefit cuts (Gauthier, 2010). Our findings suggest that a cost-benefit analysis of these policies should take into account the potential loss in mental health that would result from diminishing the comprehensiveness of maternity leave benefits.

In conclusion, we find evidence that maternity leave policies yield significant mental health benefits for working mothers, which extend beyond the period of birth and persist into older age. Our findings imply that maternity leave benefits do not only protect mothers and their children around the period of childbirth, but may contribute to healthy ageing among women during the last decades of life. This finding may have profound implications for the costs of medical care, the social participation and the productivity of older women, as well as the societal impact of older mother's mental health on family members and society.

REFERENCES

Almond, D., Hoynes, H.W., & Schanzenbach, D.W. (2011). Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes. *Review of Economics and Statistics,* 93, 387-403.

Baker, M., & Milligan, K. (2008). Maternal employment, breastfeeding, and health: evidence from maternity leave mandates. *J Health Econ,* 27, 871-887.

Berger, L.M., Hill, J., & Waldfogel, J. (2005). Maternity Leave, Early Maternal Employment and Child Health and Development in the US. *Economic Journal,* 115, F29-47.

Bitler, M.P., Gelbach, J.B., & Hoynes, H.W. (2005). Welfare Reform and Health. *Journal of Human Resources,* 40, 309-334.

Borsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., et al. (2013). Data Resource Profile: the Survey of Health, Ageing and Retirement in Europe (SHARE). *Int J Epidemiol,* 42, 992-1001.

Börsch-Supan, A., & Jürges, H. (2005). *The Survey of Health, Ageing and Retirement in Europe – Methodology.* . Mannheim: MEA.

Brockington, I. (2004). Postpartum psychiatric disorders. *The Lancet,* 363, 303-310.

Brugiavini A, Cavapozzi D, Pasini G, & Trevisan E (2013). Working life histories from SHARELIFE: a retrospective panel. *SHARE WP Series,* 11-13.

Brugiavini A, Pasini G, & Trevisan E (2012). The direct impact of maternity benefits on leave taking: Evidence from complete fertility histories. *Advances in Life Course Research,* 18, 46-67.

Brunello G, Weber G, & Weiss CT (2012). Books Are Forever: Early Life Conditions, Education and Lifetime Income. *IZA Discussion Papers 6386,* Institute for the Study of Labor (IZA).

Castro-Costa, E., Dewey, M., Stewart, R., Banerjee, S., Huppert, F., Mendonca-Lima, C., et al. (2007). Prevalence of depressive symptoms and syndromes in later life in ten European countries: the SHARE study. *Br J Psychiatry,* 191, 393-401.

Chatterji, P., & Markowitz, S. (2005). Does the Length of Maternity Leave Affect Maternal Health? *Southern Economic Journal,* 72, 16-41.

Chatterji, P., & Markowitz, S. (2012). Family leave after childbirth and the mental health of new mothers. *J Ment Health Policy Econ,* 15, 61-76.

Coser L (1974). *Greedy Institutions: Patterns of Undivided Commitment*. New York: Free Press.

Dagher, R.K., McGovern, P.M., & Dowd, B.E. (2013). Maternity Leave Duration and Postpartum Mental and Physical Health: Implications for Leave Policies. *J Health Polit Policy Law*.

Dahl, G.B., Loken, K.V., Mogstad, M., & Salvanes, K.V. (2013). What Is the Case for Paid Maternity Leave? : National Bureau of Economic Research, Inc, NBER Working Papers: 19595.

Ferrari, A.J., Charlson, F.J., Norman, R.E., Patten, S.B., Freedman, G., Murray, C.J., et al. (2013). Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med,* 10, e1001547.

Frone, M.R., Yardley, J.K., & Markel, K.S. (1997). Developing and Testing an Integrative Model of the Work–Family Interface. *Journal of Vocational Behavior,* 50, 145-167.

Gal, S., & Kligman, G. (2000). *The politics of gender after socialism: A comparative-historical essay*. Princeton: Princeton University Press.

Gauthier, A.H. (2011). Comparative family policy database, version 3 [computer file].www.demogr.mpg.de. Netherlands Interdisciplinary Demographic Institute and Max Planck Institute for Demographic Research (distributors).

Gooede, W. (1960). A Theory of Role Strain. *Am Sociol Rev,* 25, 483-496.

Gornick, J.C., & Meyers, M.K. (2003). *Families that work: Policies for reconciling parenthood and employment*. New York: Russell Sage Foundation Press.

Greenberg, P.E., Kessler, R.C., Birnbaum, H.G., Leong, S.A., Lowe, S.W., Berglund, P.A., et al. (2003). The economic burden of depression in the United States: how did it change between 1990 and 2000? *J Clin Psychiatry,* 64, 1465-1475.

Grzywacz, J.G., & Bass, B.L. (2003). Work, Family, and Mental Health: Testing Different Models of Work-Family Fit. *Journal of Marriage and Family,* 65, 248-261.

Hammen, C. (2003). Social stress and women's risk for recurrent depression. *Archives of Women's Mental Health,* 6, 9-13.

Kessler, R.C. (1997). The effects of stressful life events on depression. *Annu Rev Psychol,* 48, 191-214.

Klerman JA, & Leibowitz A. (2000). Labor Supply Effects of State Maternity Leave Legislation. In Blau FD, & E. RG (Eds.), *Gender and Family Issues in the Workplace*. New York: Russell Sage.

Klerman, J.A., & Leibowitz, A. (1999). Job continuity among new mothers. *Demography,* 36, 145-155.

Marks, S. (1977). Multiple Roles and Role Strain: Some Notes on Human Energy, Time and Commitment. *Am Sociol Rev,* 42, 921-936.

Marshall, N.L., & Barnett, R.C. (1993). Work-family strains and gains among two-earner couples. *Journal of Community Psychology,* 21, 64-78.

McCall, W.V., & Kintziger, K.W. (2013). Late Life Depression: A Global Problem with Few Resources. *Psychiatric Clinics of North America,* 36, 475-481.

Mullen, J., Kelley, E., & Kelloway, E.K. (2008). Chapter 11 - Health and Weil-Being Outcomes of the Work-Family Interface. In K. Korabik, D.S. Lero, & D.L. Whitehead (Eds.), *Handbook of Work-Family Integration* pp. 191-214). San Diego: Academic Press.

Pearlin, L.I., Schieman, S., Fazio, E.M., & Meersman, S.C. (2005). Stress, Health, and the Life Course: Some Conceptual Perspectives*. *Journal of Health and Social Behavior,* 46, 205-219.

Prince, M.J., Reischies, F., Beekman, A.T., Fuhrer, R., Jonker, C., Kivela, S.L., et al. (1999). Development of the EURO-D scale--a European, Union initiative to compare symptoms of depression in 14 European centres. *Br J Psychiatry,* 174, 330-338.

Rossin-Slater, M., Ruhm, C.J., & Waldfogel, J. (2013). The effects of California's paid family leave program on mothers' leave-taking and subsequent labor market outcomes. *J Policy Anal Manage,* 32, 224-245.

Rossin, M. (2011). The effects of maternity leave on children's birth and infant health outcomes in the United States. *J Health Econ,* 30, 221-239.

Ruhm, C.J. (1998). The Economic Consequences of Parental Leave Mandates: Lessons from Europe. *Quarterly Journal of Economics,* 113, 285-317.

Ruhm, C.J. (2000). Parental leave and child health. *J Health Econ,* 19, 931-960.

Ruhm, C.J. (2011). Policies to assist parents with young children. *Future Child,* 21, 37-68.

Ruhm, C.J., & Teague, J.L. (1998). Parental Leave Policies in Europe and North America. In M.A. Ferber (Ed.), *Women in the labour market. Volume 2* pp. 473-498): Elgar Reference Collection. International Library of Critical Writings in Economics, vol. 90. Cheltenham, U.K. and Northampton, Mass.

Schmeiser, M.D. (2009). Expanding wallets and waistlines: the impact of family income on the BMI of women and men eligible for the Earned Income Tax Credit. *Health Econ,* 18, 1277-1294.

Schröder, M. (2011). *Retrospective Data Collection in the Survey of Health, Ageing and Retirement in Europe. SHARELIFE Methodology*. Mannheim: MEA.

Sieber, S. (1974). Toward a theory of role accumulation. *Am Sociol Rev,* 39, 567-578.

Slater, P. (1963). On Social Regression. *Am Sociol Rev,* 28, 339-364.

Snyder, S.E., & Evans, W.N. (2006). The Effect of Income on Mortality: Evidence from the Social Security Notch. *Review of Economics and Statistics,* 88, 482-495.

Staehelin, K., Bertea, P.C., & Stutz, E.Z. (2007). Length of maternity leave and health of mother and child--a review. *Int J Public Health,* 52, 202-209.

Tanaka, S. (2005). Parental Leave and Child Health across OECD Countries. *Economic Journal,* 115, F7-28.

Figure Captions

Figure 1. Labor Force Participation of Women in the year of childbirth, by cohort and country, SHARE

Figure 2. Full wage weeks (maternity leave duration multiplied by % of income replaced) by country

Figure 3. Depression (Euro-D scale) by Full wage weeks of maternity leave

Figure 4. Differential depression (mothers in employment around childbirth – mothers not in employment around childbirth) by full wage weeks

Figure 5. Common trend in year of childbirth

The long-run effect of maternity leave benefits on mental health: Evidence from European countries

Research highlights

- There is controversy on whether maternity leave increases women's well-being

- This study exploits the diversity in maternity leave policies across Europe

- A comprehensive maternity leave coverage policy reduces late-life depression

- Maternity leave benefits have long-run benefits on women' mental health in older age

Supplementary Data:

The long-run effect of maternity leave benefits on mental health: Evidence from European countries

Mauricio Avendano[1,2], Lisa F. Berkman[2,3], Agar Brugiavini[4], Giacomo Pasini[4,5]

Affiliations:

1. London School of Economics and Political Science, Department of Social Policy, LSE Health and Social Care, Houghton Street, London, United Kingdom
2. Harvard School of Public Health, Department of Social and Behavioral Sciences, 677 Huntington Avenue, Boston, MA 02115, United States
3. Center for Population and Development Studies, Harvard School of Public Health, 9 Bow Street, Cambridge, USA
4. Ca' Foscari University of Venice, Department of Economics, Venezia, Italy
5. Network for Studies on Pensions, Aging and Retirement, Tilburg, The Netherlands

Corresponding Author:

Mauricio Avendano

Cowdray House, London School of Economics and Political Science, Houghton Street London WC2A 2AE

Tel.: +44 20 7955 7203

E-mail: M.Avendano-Pabon@lse.ac.uk

Table A1: : Breakdown of all samples, household response and attrition in SHARE, 2004-2008

Country	Sample size in baseline	Response rate, wave 1	Retention rate, wave 2	Retention rate, wave 3
Austria	178	56%	88%	67%
Belgium	571	39%	91%	83%
Denmark	473	63%	93%	81%
France	483	81%	93%	74%
Germany	272	63%	86%	76%
Greece	592	63%	92%	86%
Italy	543	55%	80%	85%
Netherlands	414	62%	88%	72%
Spain	337	53%	74%	79%
Sweden	361	47%	85%	75%
Switzerland	212	39%	87%	88%

Sample size refers to the baseline specification of table 2 for Austria, Belgium, Denmark, France, Germany, Italy and Spain. It refers to the robustness estimation reported in columns (19) and (2) of table A5 for Sweden, Switzerland, Greece and the Netherlands. Retention rates refer to the share of individual respondents of wave T who took part in wave T+1.

Table A2: Definitions and details of SHARE variables

	Variable description	Year of assessment
Euro-D depression scale	Standardized measure of depressive symptoms validated in the European context. The questionnaire asks respondents to report whether they had any of a list symptoms during the past month: Depression, pessimism, death wish, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment and tearfulness. A score is constructed by summing up all responses to generate a score ranging from 0 to 12.	First year of SHARE interview (either 2004/5 or 2006/7)
>3 Euro-D depressive symptoms	A measure of potentially serious depressive symptoms predicting clinical depression	First year of SHARE interview (either 2004/5 or 2006/7)
Educational level	Highest educational attainment assessed based on national levels and then reclassified into the three broad levels based on the International Standard Classification of Education (ISCED): Primary education or lower, secondary education, or post-secondary education	First year of SHARE interview (either 2004/5 or 2006/7)
Marital status	Marital status at time of first interview was assessed by asking participants whether they were married or cohabiting, separated, divorced, widowed or unmarried.	First year of SHARE interview (either 2004/5 or 2006/7)
Number of children	Number of children was assessed by asking respondents to report the birth date for all natural children, either still alive or dead.	Retrospective SHARE assessments in 2008/9
Age at childbirth	The age at which mothers had their first child, derived from the retrospective assessments in SHARE (wave 3) by combining information on the date of birth of each child and the date of birth of the mother, to obtain the age at which mothers had their first child	Retrospective SHARE assessments in 2008/9
Total tenure in the labour market	Total years women spent active in the labour market from the age of leaving full-time education or the year of their first job if they had no education. Information is derived from the retrospective assessments in SHARE, which asked individuals to report each paid job that lasted for 6 months or more including gaps between jobs, using the the life-grid History Event Calendar	Retrospective SHARE assessments in 2008/9
Log of permanent gross income	Computed according to Brunello et al (2012) and it is based on the first income reported for each job spell, plus individual labour income and income from pensions and other benefits reported in waves 1 and 2 of SHARE.	Both SHARE interviews (2004/5 or 2006/7) and retrospective SHARE assessments

The Katz Activities of Daily Living (ADL)	A measure of functional status in old age, ADLs were measured by asking respondents to report whether they had difficulties with a list of six activities to maintain basic self-care needs (bathing, dressing, toileting, transferring, continence, and eating).	First year of SHARE interview (either 2004/5 or 2006/7)
Index of Instrumental Activities of Daily Living (IADL)	A measure of functional status in old age, the IADL index assesses difficulties with relatively complex activities necessarily for independent functioning (using a map, preparing hot meals, shopping, telephone use, taking medications, housekeeping tasks, and managing money)	First year of SHARE interview (either 2004/5 or 2006/7)
GALI activity limitation index item	The GALI (Global Activity Limitation Index) item asked participants to report to what extent during the past six months at least, they were limited because of a health problem in activities they usually do. Response included severely limited, limited but not severely, or not limited. Responses were dichotomized based on no limitations vs. any limitations.	First year of SHARE interview (either 2004/5 or 2006/7)
Years smoking	The years individuals smoked over the entire life was derived from information on the age at which individuals first smoked; whether they smoked at the present time; and the age at which they stopped smoking if they had stopped. The measure does not include short smoking interruptions within the period the individuals was primarily a smoker.	First year of SHARE interview (either 2004/5 or 2006/7)
Days drinking in a month	Respondents are asked to report how often they drank alcoholic beverages, like beer, cider, wine, spirits or cocktails in the three months prior to the interview (almost every day, 5-6 time a week, 3-4 times a week, 1-2 times a week, 1-2 times a month, less than once a month, never)	First year of SHARE interview (either 2004/5 or 2006/7)
Number of Miscarriages	The number of miscarriages is retrieved from the retrospective SHARE questionnaire where respondents are asked to report the year in which they had each miscarriage.	Retrospective SHARE assessments in 2008/9

Table A3: Descriptives of maternity leave policy database: Weeks of maternity leave, percentage of wages covered during leave and full wage weeks per country and period (1960-1979 and 1980-2010)

Country	Period	Weeks of paid maternity leave			Percentage of wage (in manufacturing sector) covered during leave			Weeks of Full wage leave*		
		Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
Sweden	*1960-1979*	12.8	39.0	25.3	45.0	90.0	63.8	5.8	35.1	17.2
	1980-2010	39.0	68.6	59.3	62.0	90.0	72.9	27.7	57.6	42.8
Denmark	*1960-1980*	14.0	14.0	14.0	19.0	90.0	59.3	2.7	12.6	8.3
	1980-2011	18.0	18.0	18.0	39.0	90.0	62.2	7.0	16.2	11.2
Germany	*1960-1981*	12.0	14.0	13.0	100.0	100.0	100.0	12.0	14.0	13.0
	1980-2012	14.0	14.0	14.0	100.0	100.0	100.0	14.0	14.0	14.0
Netherlands	*1960-1982*	12.0	12.0	12.0	100.0	100.0	100.0	12.0	12.0	12.0
	1980-2013	12.0	16.0	14.3	100.0	100.0	100.0	12.0	16.0	14.3
Belgium	*1960-1983*	12.0	14.0	13.0	60.0	79.5	64.9	7.2	11.1	8.5
	1980-2014	14.0	15.0	14.7	76.4	79.5	77.4	10.7	11.5	11.3
France	*1960-1984*	14.0	16.0	14.1	50.0	90.0	68.0	7.0	14.4	9.6
	1980-2015	16.0	16.0	16.0	84.0	100.0	93.3	13.4	16.0	14.9
Austria	*1960-1985*	12.0	16.0	13.2	100.0	100.0	100.0	12.0	16.0	13.2
	1980-2016	16.0	16.0	16.0	100.0	100.0	100.0	16.0	16.0	16.0
Italy	*1960-1986*	17.0	23.7	21.3	80.0	80.0	80.0	13.6	18.9	17.1
	1980-2017	21.5	23.7	21.9	80.0	82.0	80.1	17.2	18.9	17.5
Spain	*1960-1987*	12.0	14.0	12.3	60.0	75.0	69.7	7.2	10.5	8.6
	1980-2018	14.0	16.0	15.3	75.0	100.0	88.7	10.5	16.0	13.7

**Full wage weeks are the product of weeks of paid maternity leave and percentage of wage (in manufacturing sector) covered during leave*

Table A4: Number of first maternity episodes occurring between age of 16 and 25, by country of residence of the mother and FWW in place in the specific country and year of childbirth

Full wage weeks	AT	DE	ES	IT	FR	DK	BE	
3	0	0	0	0	0	68	0	68
4	0	0	0	0	0	64	0	64
5	0	0	0	0	0	47	0	47
7	0	0	105	0	258	0	253	616
8	0	0	0	0	0	29	162	191
9	0	0	201	0	0	0	0	201
11	0	0	70	0	0	25	177	272
12	158	167	0	0	0	45	0	370
13	0	0	0	0	206	164	0	370
14	0	120	0	97	57	0	0	274
16	41	0	1	0	0	46	0	88
17	0	0	0	284	0	0	0	284
19	0	0	0	181	0	0	0	181
Total	199	287	377	562	521	488	592	3026

Table A5: additional robustness checks

	(1)	(2)	(3)	(4)	(5)
Employed around childbirth	0.174	0.140*	0.041	0.042	1.084
	(0.109)	(0.083)	(0.027)	(0.029)	(0.063)
Comprehensive maternity leave	0.352**	0.356**	0.070*	0.077*	1.151*
	(0.171)	(0.162)	(0.038)	(0.041)	(0.091)
Employed around childbirth*comprehensive maternity leave	-0.366**	-0.348**	-0.097**	-0.100**	0.819**
	(0.161)	(0.145)	(0.038)	(0.040)	(0.065)
Age	0.083	0.195	0.028	0.03	1.091
	(0.183)	(0.140)	(0.039)	(0.044)	(0.1)
Age squared	-0.001	-0.002	0	0	0.999
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
High school education	-0.149	-0.175**	-0.015	-0.017	0.971
	(0.096)	(0.080)	(0.023)	(0.024)	(0.046)
College education	-0.122	-0.203**	-0.018	-0.02	0.959
	(0.121)	(0.103)	(0.031)	(0.034)	(0.068)
Married/Cohabitation or Single	-0.464***	-0.449***	-0.124***	-0.138***	0.791***
	(0.171)	(0.146)	(0.046)	(0.050)	(0.063)
Number of children	0.063*	0.077**	0	0.001	0.998***
	(0.035)	(0.031)	(0.008)	(0.008)	(0.014)
1931-1940 cohort	-0.66	.	-0.554***	.	0.369**
	(0.856)	.	(0.183)	.	(0.161)
1941-1950 cohort	-0.641	-0.022	-0.580***	-0.027	0.345**
	(0.881)	(0.160)	(0.188)	(0.052)	(0.156)
1951-1960 cohort	-0.29	0.288	-0.544***	0.006	0.376**
	(0.798)	(0.227)	(0.171)	(0.072)	(0.156)
1961-1970 cohort	-0.235	0.573	-0.501***	0.042	0.409***
	(0.557)	(0.442)	(0.123)	(0.135)	(0.131)
1971-1980 cohort	.	1.065	.	.	.
	.	(0.691)	.	.	.
Activities of daily living (ADL) score	0.361***	0.299***	0.009	0.011	1.002
	(0.104)	(0.096)	(0.023)	(0.033)	(0.031)
Instrumental activities of daily living (IADL) score	0.551***	0.484***	0.084***	0.107***	1.122***
	(0.096)	(0.082)	(0.016)	(0.025)	(0.025)
Limitations with activities (GALI)	1.115***	1.079***	0.192***	0.197***	1.508***
	(0.086)	(0.074)	(0.020)	(0.021)	(0.062)
Years smoking	0.005*	0.005**	0	0	1.001
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)
Days drinking in a month	-0.001	-0.001	0	0	1.000
	(0.004)	(0.003)	(0.001)	(0.001)	(0.002)
Year of child birth	-0.098***	-0.076***	-0.023***	-0.026***	0.947***
	(0.027)	(0.025)	(0.007)	(0.007)	(0.014)
N obs	3218	4436	2881	2880	2880

Notes: Stars represent statistical significance: * p<0.1. ** p<0.05. *** p<0.01. Standard errors reported in parenthesis are robust to heterskedasticity. Excluded cohort is 1921-1930. Germany is the excluded country specific trend in year of childbirth. Comprehensive maternity leave takes value 1 if the value of FWW (full wage weeks) is larger than 12 weeks in Denmark, Germany, France and Austria; larger than 16 weeks in Italy; and larger than 8 weeks in Belgium and Spain. Column (1) includes all women aged 16-25 at childbirth and adds Sweden, where comprehensive maternity leave takes value 1 if FWW is larger than 16. Column (2) adds Greek, Swiss and Dutch women aged 16-25 at childbirth in the control group (all those countries did not experience any policy change in the sampled period in which women faced low fww policies). The

dependent variable in columns (3), (4) and (5) is a dummy variable which takes value 1 if eurod is larger than 3, a threshold commonly used for depression caseness, and 0 otherwise. All columns include the baseline countries and define the treatment and control group according to the baseline definitions (see footnote to table 2). Column (3) reports coefficients from a Linear Probability model, column (4) marginal effects from a probit regression; column (5) risk ratios from a Poisson regression.

Table A6: Mean and standard deviation (sd) of dependent variable and controls by treatment vs control group

	Treatment group		Control group		test of equality	
	Mean	Sd	Mean	Sd	t-test	p-value
Euro-D depression scale	2.621	2.241	2.991	2.479	4.084	<0.0001
Age at interview	60.418	5.998	61.348	6.562	3.872	0.0001
Elementary education	0.300	0.458	0.392	0.488	5.128	<0.0001
High school education	0.475	0.500	0.333	0.472	-7.728	<0.0001
College education	0.200	0.400	0.115	0.320	-6.287	<0.0001
Married/Cohabiting	0.947	0.225	0.976	0.154	4.076	<0.0001
Number of children	2.358	1.059	2.822	1.409	9.598	<0.0001
1931-1940 cohort	0.120	0.325	0.173	0.378	3.892	0.0001
1941-1950 cohort	0.494	0.500	0.500	0.500	0.300	0.764
1951-1960 cohort	0.370	0.483	0.304	0.460	-3.691	0.0002
1961-1970 cohort	0.016	0.126	0.023	0.149	1.232	0.218
1971-1980 cohort	0.000	0.000	0.001	0.029	1.000	0.318
Activities of daily living (ADL)	0.083	0.421	0.121	0.529	2.051	0.04
Instrumental ADL's (IADL)	0.150	0.533	0.222	0.711	2.945	0.003
Limitations with activities	0.354	0.478	0.418	0.493	3.461	0.001
Years smoking	10.044	14.528	7.438	13.543	-4.918	<0.0001
Days drinking in a month	8.552	10.609	6.786	10.630	-4.381	<0.0001
year of childbirth	1970.650	6.080	1969.347	6.302	-5.528	<0.0001

The NIHMS has received the file 'Maternity et al Supplementary data R3.docx' as supplementary data. The file will not appear in this PDF Receipt, but it will be linked to the web version of your manuscript.