Women's employment, income and divorce in West Germany: a causal approach

Daniel Brüggmann[*]

Abstract

In this paper, I assess the employment and income effect of divorce for women in West Germany between 2000 and 2005. With newly available administrative data that allows me to adopt a causal approach, I find strong negative employment effects with respect to marginal employment and strong positive effects with respect to regular employment. However, in sum, the overall employment rate (marginal and regular employment combined) is not affected. Furthermore, the lower the labor market attachment before separation is, the more pronounced employment effects are. In addition, I also estimate the impact of divorce on daily gross incomes. I find no convincing evidence for an income effect. I conclude that a divorce might have a pure labor supply effect only.

Keywords: Divorce, Female employment, Propensity score matching, Difference-in-difference, Boosted regression, Principal stratification, Sample selection

JEL Classification: C14, J3, J12, J22

1 Introduction

Divorce and separation rates have increased in most industrialized societies since the 1960s. In the European Union, for example, the crude divorce rate stood at 0.8 in 1965. This figure soared to 1.5 in 1980, to 1.8 in 2000 and 1.9 in 2015 (Eurostat 2018). In response to this development, a large body of work has amassed that examines the impact of separation or divorce on either economic well-being or on changes in labor market activities (Hauser et al. 2016; Bröckel and Andreß 2015; Tamborini et al. 2015; DiPrete and McManus 2000; van Damme et al. 2009; Jenkins 2008; Mueller 2005; Raz-Yurovich 2011; Tach and Eads 2015; McKeever and Wolfinger 2001). Research by Hauser et al. (2016) for Germany for the period between 1990 and 2006 has shown that women experience a dramatic short-term drop in equalized household income of approximately 26% in the year following the dissolution of a marital or cohabiting union. Government taxes and transfers reduce this decline to 17%. While there is a significant drop for women, the

equalized household income before taxes and transfers of men increases by 4% after separation and it only drops by 4% from the pre-divorce income once taxes and transfers are taken into account (Hauser et al. 2016).

In this paper, I add to the previous literature by using administrative data to examine the causal consequence of divorce on individual labor income and employment participation of women in West Germany. Previous research for Germany was regularly constrained by the low number of events available in social science surveys that were used to study the economic ramifications of divorce and separation. Thus, scholars often combined multiple survey years or even decades for their investigations (Hauser et al. 2016; Bröckel and Andreß 2015; DiPrete and McManus 2000). In this paper, I overcome some of these limitations by focusing the analysis on women with a divorce file opening in the calendar year 2002 using administrative data of the German pension insurance. Apart from the overall employment rate (which is defined as being marginally and/or regularly employed) and the rate for regular employment, I also examine changes in marginal employment. In the context of the German system, a transition from marginal to regular employment is

*Correspondence: d.brueggmann@phd.hertie-school.org
Hertie School, Friedrichstrasse 180, 10117 Berlin, Germany

a significant process. Marginally employed persons face lower wages, are exempt from unemployment benefits, do not contribute to the statutory health insurance and, until 2013, were only voluntarily covered in the statutory pension system. As many married women are working marginally in Germany, it is important to understand whether divorce increases regular employment.

As a method, I primarily rely on propensity score matching (kernel matching). Matching techniques have become widely used to unravel causal effects. In a setting like divorce where the selection into divorce is not random, the "divorce effects" in conventional models are very likely biased. The matching approach is one possibility to address the selection bias. It removes selection into divorce by finding similar individuals in the treatment and control group (conditional on observed pre-treatment characteristics). Thus, based on observed covariates it mimics a randomized controlled trial.

As to the structure of the analysis, I first examine the employment effects for marginal[1] employment, regular[2] employment and then I estimate the overall employment rate as a combination of both. Since the plausibility of estimates relies heavily on the assumption of conditional independence (no hidden bias), I scrutinize the employment effects with respect to hidden bias from unobserved confounders (Mantel–Haenszel bounds) (Mantel and Haenszel 1959). In a second step I analyze the impact of divorce on daily gross earnings (for regular employment only) by principal stratification (Zhang and Rubin 2003; Zhang et al. 2008; Lee 2009; Huber and Mellace 2015). I decided on principal stratification because in the presence of sample selection (non-random selection into employment) naïve treatment-minus-control differences cannot be interpreted as impact estimates (Lee 2009).

2 Institutional background

For a long time, women in West Germany were treated primarily as housewives and caregivers instead of workers or breadwinners and various institutional features fostered the gendered or traditional division of labor between spouses.

In particular, the tax-splitting scheme provides strong incentives for both spouses to combine one large labor income with one small or zero labor income. The splitting advantage was as high as € 8000 for high earner breadwinners and was close to € 3000 for an average breadwinner (Steiner and Wrohlich 2004). This tax advantage

strongly inhibits women's labor market participation due to the relatively high marginal tax rate for the "second earner". If the wife were to increase her labor income, the splitting advantage would be reduced with each Euro additionally earned until both spouses earn the same.

Apart from the tax system, availability of childcare influenced parents' ability to participate in the labor market (Uunk 2004). Childcare provision has increased over time in West Germany, but public childcare was largely restricted to part-time care for children of pre-school age (age 3–6) (Wrohlich and Müller 2014). Since 2005, the German government has initiated several reforms to increase the provision of day care for children under age three. However, for the period that I investigate (2000 to 2005), availability of full-time care and day care for children under age three was very restricted (Bröckel and Andreß 2015). In addition, the long duration of parental leave was considered an obstacle for women's swift return into the labor market after childbirth and the low amount of benefits was regarded as a barrier for fathers' uptake (Spieß 2011). It was only in 2007 that the German government initiated a major reform and introduced an income-related "*Elterngeld*". This reform is, however, not relevant for my investigation as it was enacted after the observation period.

As for divorce regulations, until 2008 German law offered the possibility of receiving support payments for the economically weaker spouse (§1361 BGB) and the amount of alimony was granted based on the living conditions before divorce. The lower earning partner (usually the woman) was, in addition, not expected to take up employment until the child entered primary school, and was not expected to work full-time before the youngest child reached age 16 (Bröckel and Andreß 2015; Hummelsheim 2009).

While family policies did not see significant shifts around 2002, there have been major labor market reforms since 2003, including the Hartz reforms. While the Hartz IV reform in 2005 involved a drastic cut in benefits for the long-term unemployed and stricter job search obligations, the Hartz II reform in 2003 provided incentives for the uptake of marginal employment by lifting the maximum income from € 325 to € 400 and exempting marginal employment (held as a secondary job) from social security contributions. In theory, the latter reform could partly affect my estimates and result in overestimating the true unbiased treatment effect of divorce as long as married women react stronger to the incentives than divorced women. With the approach applied here, I was not able to disentangle the reform effect from the divorce effect. However, the comparisons of treatment effects for marginal employment before the reform (2002), at the reform year (2003) and after the reform (2004 and 2005)

---

[1] Marginal employment (or equally called mini-jobs in Germany) are specific employment types with an earnings threshold of € 400 in 2003.

[2] Regular employment refers to standard employment contracts for full-time or part-time jobs with social security contributions.

show no strong deviations. I conclude that the likelihood of deviations from the true unbiased effect is rather low.

Overall, social policies in Germany supported, until very recently, the male breadwinner model where one partner reduced employment while married. Despite an increase in women's employment rate over time, the large majority of women (especially with children) did not work full-time, but were employed part-time or marginally (Bröckel and Andreß 2015; Engstler and Menning 2004). Especially marginal employment is widely considered as ambivalent because being continuously employed in the marginal sector means a prolonged risk of dequalification, wages at the lower end with limited access to in-house training and career advancement (Seifert 2011). However, compared to non-employment, marginal employment might ameliorate the depreciation of human capital and serves as a stepping-stone into regular employment if employers use it as a screening mechanism (Caliendo et al. 2012).

3 Prior findings

A large body of literature has amassed that studies the social and economic consequences of separation and divorce on equivalent household income. In most instances these studies have found substantial declines (before and after government taxes and transfers) for separated women in the US (Hauser et al. 2016; Tach and Eads 2015; McKeever and Wolfinger 2001; DiPrete and McManus 2000), in Europe (Uunk 2004), in the UK (Jenkins 2008) and in Germany (Hauser et al. 2016; Bröckel and Andreß 2015; Burkhauser et al. 1991). While the majority of empirical assessments have addressed changes in household income, others have investigated the effect of divorce on women's employment and earnings. Studies on the employment effect mostly show that women increased their labor supply after break-up. Raz-Yurovich (2011) analyzed the Israeli context, for example, and found that women increased their employment stability and the number of jobs held following divorce. Monthly salary increased only slightly and the effect was not significant.

Tamborini et al. (2015) studied women's employment and average earnings in 1970–1974, 1980–1984 and 1990–1994 in the US. They found long-lasting employment and income increases. However, employment and income increases were substantially lower in the latter period. The decline in effect size is explained by the increased labor market activity while married because women who are already more involved in the labor market may be limited in how much they increase their employment.

While most studies found that divorce leads to an increase in women's employment, there are also studies finding the opposite (Mueller (2005) for Canada, Jenkins (2008) for the UK and Van Damme et al. (2009) for countries in Europe). Jenkins (2008), for example, found lower employment rates after divorce in the UK. In the period 1991–1997, employment dropped by 5 percentage points (pp) and in 1998–2003 by 2 pp. The most obvious reason for the lower drop in the second period were policy changes in 1998, which increased the incentives to work.

Van Damme et al. (2009) studied the employment effects in Europe for 13 countries in the period 1994–2001. They found a significant but small increase in participation rates after divorce. Overall, the increase was from 63% the year before separation to 68% 1 year after, but country variations were substantial. While in the Netherlands, Denmark and Italy the increase was more than 10 pp, negative but not significant results were found for Finland and Greece. Employment in the UK dropped significantly by 4.9 pp. Overall, increases in employment were greatest for those countries where women worked less before divorce. For Germany, where female employment rates are low, they found an overall increase of 7.3 pp to 76%.

The German context was analyzed for example by Hauser et al. (2016) and Bröckel and Andreß (2015) based on before-after estimations. On average, divorced women in West Germany increased their employment rates by 8 pp to 74% in the period 1990–2006 (Hauser et al. 2016) and by 6 pp to 73% in 2000–2012 (Bröckel and Andreß 2015). Average labor earnings (of those who were employed) increased by 36% to € 17,775 and by 22% to € 14,681. In contrast, DiPrete and McManus (2000) found for the period 1984–1996 (based on a fixed-effect approach predicting the 2-year change around union dissolution) a slight, non-significant negative impact of divorce on labor earnings.

I contribute to the existing literature in the following way. I estimate the "treatment effect" of divorce on the employment rate and on daily gross incomes. This means that I compare divorce effects to a well-defined control group. While for the employment rate the treatment-minus-control difference can be a valid estimate (if matching successfully randomizes the divorce status like random assignment would do), the analysis of incomes, however, might still be flawed. The reason is that earnings are only observed conditional on being employed. As Lee (2009) notes, even with the aid of a randomized experiment, the analysis of an outcome (income) which is dependent on another outcome (employment) is subject to the sample selection problem, if the first outcome (employment) is not randomly distributed after the impact of the treatment. It seems very plausible that for some women (i.e. those women with no children, with older children and women with better education)

employment is easier to find. Thus, employment after the treatment is not random but a matter of children and education. Likewise, those women might also work more hours and thus, have higher daily earnings. Therefore, the simple gross daily income comparison between treated and controls might be flawed by the characteristics that promote employment. To overcome this shortcoming in my analysis, I use the principal stratification framework (Zhang and Rubin 2003; Zhang et al. 2008; Lee 2009; Huber and Mellace 2015). To my knowledge, I am the first who applies this concept to the divorce literature.

4 Theoretical considerations and key questions

Prior evidence has shown that employment effects vary by countries and time periods. It has also been shown that divorce may cause an increase or a drop in labor market participation. There are arguments for both effect directions.

On the one hand, the loss in economies of scale as well as the shock in household income should, ceteris paribus, increase financial pressure and reduce the reservation wage. One might also argue that the family is maximizing a joint family utility function (Killingsworth and Heckman 1987) or is specializing in home and labor work (Becker 1981) while married. As new information becomes available and marriage quality decreases, the value of specialization and the value of maximizing a joint family-utility might change and the focus turns to individual utility and the importance of women's loss in labor market skills. This again reduces the reservation wage because women gain from increasing their work effort in order to acquire work experience for the purpose of employability and income prospects after separation.

On the other hand, since divorcees might face time constraints (especially mothers), qualify for welfare payments or maintenance payments, or move into smaller homes, the reservation wage might also be unaffected or may even increase if women adapt to the new economic condition of reduced household income. Moreover, even if women (in particular mothers with young children) would like to work, there remains the obstacle of low public childcare availability for children under age 3. Although childcare availability has increased over time in West Germany, the share of children under 3 in day care was only 7.7% in 2005 (Bröckel and Andreß 2015). Therefore, the non-availability of public childcare very likely hampered mothers' labor market entry.

Summing up, a theoretical assessment of the overall effect of divorce on employment is ambiguous. However, one can expect strong effect heterogeneity by whether the woman had been attached to the labor market prior to divorce. Women who were only working in marginal employment should face strong economic incentives to expand their labor market attachment by shifting to regular employment. Conversely, regularly employed women and those with a strong labor market attachment before separation will not expand their employment to the same degree. Contrary, they might need to decrease it if the double burden of employment and childrearing increases.

Besides employment effects, I also study the impact of divorce on daily gross earnings. In contrast to married women, divorced women might be in need to upwardly adjust their daily income because financial strains are higher and household income is lower (to the extent that alimony and governmental payments are not counteracting those adjustments). On the other hand, due to the double burden of employment and childrearing (in the case of mothers) divorcees might be less able to participate in on-the-job training and might even be forced to change jobs to mother-friendly jobs and to trade higher earnings for flexibility (Gangl and Ziefle 2009).

5 Data and method
5.1 Data

In the present study, I used administrative data from the statutory German pension system. I linked the records of the Sample of Active Pension Accounts (VSKT) with the records of the Pension Rights Adjustments Statistic (VA). The VSKT is a one percent random sample of all individuals with a pension account in Germany. It provides detailed pension-relevant information, such as information on the individuals' employment and earnings history, spells of parental leave, and childbirths since age 15 (Stegmann and Himmelreicher 2008). The VA contains the dates of separation and divorce of those individuals who have gotten divorced since 1977 and whose pension entitlements were equalized after divorce. The pension fund collects these data, because Germany has a system of "income splitting", whereby pension entitlements are split after divorce (for more details, see Keck et al. 2019). The great advantage of using these data is first, that they provide a reasonably large sample size for a divorce event in a single year and second, the high accuracy of the data (because these data is the source for pension calculations). Furthermore, unlike prospective survey data, administrative data do not suffer from attrition, which is especially likely to occur after a separation or a divorce. However, there are other caveats that I need to mention. One limitation of the data is that the administrative data (the source data for the VSKT) do not include the full resident population, but cover only those who have a pension account. About 90% of the resident population in Germany are included in the data, but people in certain professions, such as civil servants and farmers, are

not included (Kruse 2007).[3] Furthermore, not all divorces are included in the VA because the data only contain information on divorces that result in pension splitting. Pension splitting is, in theory, mandatory, but certain couples—and particularly those with short marriages—can avoid pension splitting (Keck et al. 2019). Thus, the observed divorcees might not be a representative sub-population of all divorcees in Germany. This would limit the external validity of the study. For that reason, my results are limited to the population of women with pension right adjustments in the divorce. However, note that about two thirds of all divorces are included in the data (Keck et al. 2019).

I have restricted the sample to persons with a divorce file opening in 2002. I have further restricted the sample to women who were 25 to 55 years old, were married at least 5 years before the file was opened, are of German citizenship and lived in West Germany (i.e. never earned any pension records in East Germany). The final analytical sample consist of 413 divorced women. Note that I dropped East German women from the analysis first, due to low case numbers, second, because of structural differences in childcare availability between West and East Germany and lastly, because of systematic differences between West and East German women in terms of labor market participation.

Separation (t_0) is defined as the 15th day of the month in 2002 that the divorce file was opened; i.e., the month when the defendant received the divorce petition. I have furthermore limited the investigation to the time window of 2189 days before the file was opened up to 1095 days thereafter. Employment and income effects are then estimated at file opening (t_0), 1 year after (t_{365}), 2 years after (t_{730}) and 3 years after (t_{1095}).

For my control group I used married women out of the same combined dataset who were still married in 2002 but experienced a divorce in the distant future (after 2008). Taking the women from the same dataset had the advantage that I indirectly controlled for variables that I usually cannot observe (like preferences to work, motivation or religiosity) but which are important for the selection into divorce and employment. To the extent that a woman who is married and who never gets divorced faces lower divorce risks, lower employment risk and follows more closely traditional family norms, my results would be upward biased if these women were chosen as the control group. A control group instead who eventually shares the same risk to divorce controls for

such unobserved characteristics and reduces the risk of overestimation.

In total, the control group consists of 1437 women who fit [at a randomly chosen month (15th day) in 2002] the same criteria as the treatment group except that they had no file opening in 2002. The control group consisted of 262 women with a file opening in 2008, 267 women with an opening in 2009, 219 women in 2010, 208 in 2011, 176 in 2012, 160 in 2013, 117 in 2014 and 28 in 2015.

I also split the main sample into four subsamples in order to derive employment and income effects for women with different labor market attachment while married. The subsamples were constructed first, by cumulating the days of regular employment between t_{-2189} to t_{-730} and second, by generating four quantiles.[4] However, I display results only for the most extreme groups, i.e. the subsample of women with 0 days of regular employment between t_{-2189} and t_{-730} (Group A; $N_{treated} = 144$ and $N_{control} = 654$) and the group of women with strong labor market attachment, i.e. days ≥ 967 (Group B; $N_{treated} = 134$ and $N_{control} = 328$).[5] I focused on these subgroups because each presents an extreme part of women's labor market attachment while married, i.e. they represent the typical housewife or mother on one side with relatively low lifetime work commitment and, on the other side, the women with substantially more work commitment and fewer young children (see Tables 6 and 7, Appendix for selected demographic statistics).

A practical challenge is the causal direction of female labor supply and divorce, and addressing the competing perspectives, i.e. the "anticipation" or the "independence" perspective [for a detailed discussion see Özcan and Breen (2012)]. I followed the practice in prior studies and implied anticipation of a divorce, i.e. all employment and income changes refer to the baseline day at t_{-730} instead of t_0. However, I also addressed the independence perspective by the framework of matching and the chosen pre-treatment period (t_{-2189} to t_{-730}). Thus, I controlled for observed differences between divorcees and married women in the period t_{-2189} to t_{-730} (except childbirth).[6] In addition, since higher education, occupational training and work experience are important determinants for employability, income prospects and marital stability

---

[3] Some occupations are not fully covered by the German pension system because those occupations have their own pension institutions and are not obliged to contribute to the statutory pension system. Those occupations are for example architects, medics or self-employed individuals.

[4] The cumulated days for regular employment within t_{-2189} and t_{-730} are 0 days for the first group, are 2 to 129 days for the second group, are 131 to 960 days for the third and 967 to 1461 days for the fourth group.

[5] Case numbers for the second quantile are $N_{treated} = 29$ and $N_{control} = 96$ and for the third quantile $N_{treated} = 103$ and $N_{control} = 359$.

[6] I measure childbirth in the period t_{-729} to t_{-365}, since childbirth occurs with a time-lag of 9 months and the decision to become pregnant often lies well before t_{-730}. Note, marginal employment is not recorded before 1998, thus, t_{-2189} to t_{-1825} and t_{-1824} to t_{-1460} are excluded for marginal employment and income measures.

(following the independence perspective) I also constructed lifetime measures. These measures are cumulated days for the entire period of age 15 to t_{-730}. A full list of all covariates is presented in Table 5 (Appendix).

5.2 Method

The abovementioned covariates (Table 5, Appendix) were used in linear form in a logit regression to estimate the individual probability for a file opening in 2002. This is the propensity score.[7] In addition, I used a second model from machine learning as an alternative way to calculate the propensity score. This model is based on random trees and incorporates many higher order and interaction terms and thus acknowledges that the true functional form of the selection process was unknown. I used a general boosted model (GBM) for three reasons: first, because these models can handle large numbers of covariates, second, these models are immune to multicollinearity and third, because they often achieve better balance properties than simple logistic regressions (McCaffrey et al. 2013).[8]

Because estimated propensity scores are highly sensitive to selected covariates and their interactions I expect strong differences between these two models. However, if both models come to similar point estimates for employment effects (regardless of strong differences in estimated propensity scores) I am confident that the model is robust against misspecification.[9]

These estimated propensity scores were used to derive weights by either kernel matching or weighting by the odds. To be precise, I combined the logit model with kernel matching and the GBM model with weighting by the odds.[10]

Based on these derived weights, I estimated the average treatment effect on the treated (ATT), i.e. I estimated the effect of divorce on employment for those women with a file opening in 2002. In this set-up, the control group serves as a reflection of the outcome that the treated group would have experienced had they not filed for divorce. For my purpose, I combined matching with a difference-in-difference (DiD) approach, thereby considering the change in employment from the baseline day t_{-730} to the respective day at either t_0, t_{365}, t_{730} and t_{1095}.[11]

The mean values of the outcome variable of the control group only serve as a reflection of the outcome that the treated group would have experienced had they not filed for a divorce, if the following assumptions are satisfied: Stable Unit Treatment Value Assumption (SUTVA), Conditional Independence Assumption (CIA) and common support.

The SUTVA assumption rules out that the treatment affects the control group, i.e. we need to assume that the job search effort of the divorcees does not affect the employment probability of married women. Otherwise, the outcome of the control women would not be the same as the one they would have experienced in a world without divorcees and the counterfactual outcome would be biased, leading to overestimated results. Since I have only microdata I am not able to estimate such displacement effects on the macro-level and, thus, I am not able to verify that such effects do not exist. However, I assume that the labor market in Germany is large enough and can absorb all women (from the treatment and from the control group) without placing constraints on one group. This assumption might be reasonable because first, the entry into divorce is quite low in comparison to the number in unemployment. Second, a substantial part of divorcees is already employed while married and third, divorcees might aim for regular employment whereas married women are often marginal employed and stay marginal employed (thus, competition for the same jobs might be rather low).

It is in general difficult to claim that the CIA holds because it rules out the existence of unobserved covariates that simultaneously affect treatment and employment decisions. I therefore addressed this issue separately in the sensitivity analysis by scrutinizing the employment effect with respect to hidden bias from unobserved covariates.

---

[7] The propensity score estimation is used only as a tool to get covariates balanced. The concern is not about the parameter estimation of the covariates, but the resulting balance property and thus, standard concerns about collinearity do not apply (Stuart 2010).

[8] In particular, I use *TWANG* from the *R* library with the following parameters: Interaction depth (3); smoothing parameter (0.0001); iterations (1,000,000) and stopping rule (minimizing ND_{mean}).

[9] Both models (logit with linear covariates and GBM with higher order covariates) come to very different propensity scores. The largest observed difference is 0.44 probability points (for one woman the logit-based propensity score is 0.70 and it is 0.26 for the same woman in the GBM model).

[10] I extract the kernel weights from kernel matching (Epanechnikov) with *PSMATCH2* at a bandwidth $h=0.056$ for my main sample and 0.082 for group A and 0.038 for group B. Odd weights are derived by $w_{i,j} = D_i + (1 - D_j)\frac{ps_j}{1-ps_j}$ with $D \in (0,1)$ if treated or not and ps as the propensity score. Subscripts i for treated and j for control. Extreme weights can be a problem for odd weighting (if women from the control group have high propensity scores) because results are dominated by only a few cases. In my study, however, odd weights range between 0.033 and 1.18 and a mean of 0.42. The distribution of weights is therefore reasonable without extreme outliers.

[11] Note, I applied DiD as a procedure to remove any pretreatment differences in the outcome of interest after matching, i.e. to remove the difference in outcome between treated and control group at t_{-730} from the simple ATT (i.e. the difference in outcome between treated and control at t_0, t_{365}, t_{730} and t_{1095}). In other words, I did not rely on the common trend assumption for the identification of the treatment effect. Lechner (2011) showed that DiD and matching assumptions do not nest in each other and that the researcher has to decide on which identifying assumptions the analysis is based, i.e. either DiD assumptions or matching assumptions but not both. I relied on the matching assumptions.

Lastly, since I applied kernel matching with reasonably small bandwidths, I claim that the common support assumption is fulfilled automatically.

5.3 Summary statistics

In Table 1, I compare my treatment and control group on some selected background characteristics (for subsamples see also Tables 6 and 7, Appendix). The raw sample (column 1 and 2) shows that the characteristics of the women who did not undergo a divorce differed sharply from the characteristics of the divorcees. The most obvious differences are found in age, in marriage duration, in childbirth and the number of children, and the labor market outcomes of regular employment.[12] Divorcees are on average older at t_0, have been married longer, are less likely to have younger children, are more often regularly employed and have higher incomes (income ≥ 0). The low share of young children under six in the treated group might be a sign that young children reduce the risk of divorce or that opportunity costs of divorce are higher. A more formal analysis of the selection process for the main sample (before and after matching) is shown in Appendix (Table 5, column 1 and 2).

After matching (Table 1, columns 4, 5 and 6) both groups are rather similar and the difference between the treated and matched married women is almost eliminated. The largest difference is in *days of work disability* with 8% of a standard deviation (column 6). The value, nevertheless, is low and does not show a serious bias. Following Sianesi (2004), the matching procedure succeeded in eliminating observed differences between treated and controls, as indicated by the low Pseudo R^2 of 0.003 after matching (Table 5, column 2, last row, Appendix).[13][14]

6 Results

6.1 Empirical findings—employment dynamics

For my analysis, I estimated the change in overall (i.e. marginal and/or regular), marginal and regular employment for the day of the file opening, 1, 2 and 3 years after the file opening (t_0, t_{365}, t_{730}, t_{1095}) to the baseline day at t_{-730}. The difference in the change (DiD) between the treated and the controls shows the effect of divorce for

---

[12] The justification whether mean values differ is based on the Normalized Difference known from Rosenbaum and Rubin (1985).

[13] Note, Pseudo R^2 reduction due to matching are similar for group A (Pseudo R^2 reduced from 0.0673 to 0.0078) and group B (from 0.1383 to 0.0107).

[14] Table 8 (Appendix) provides additional balance statistics for the subgroups. The test statistics (Normalized Difference and Kolmogorov–Smirnov) show no strong deviation from randomizing individuals into treated and control group for kernel matching. The GBM model (with odd weighting), however, performed more poorly but balance results are still sufficient and reliable.

those women with a divorce file opening in 2002. To the extent that the CIA is satisfied, the outcome of the control group would be the outcome that the treated group would have experienced had they not divorced. For the moment, I assume that the CIA holds and assume that selection on unobservable confounders is irrelevant.

In the main sample (Table 2, panel 1), the overall divorce effect is significant and -9 percentage points (pp) for marginal employment and 8 pp for regular employment in t_0, i.e. marginal employment is 9 pp lower and regular employment is 8 pp higher than it would be without divorce. The effect on the overall employment rate is not significant, slightly decreases and shows that it might not be the best parameter to look at because important changes in employment types are hidden.

Figures 1 and 2 visualize the employment rates for treated and controls and show that the change in employment rates in marginal and regular employment is driven by the employment dynamic of the divorcees but not by the married women. While the labor market participation of women from the control group is fairly stable over time, I observe signs of anticipation in the treatment group, starting around 1 year before the divorce file was opened (Figures 6, 7, 8 in Appendix provide the effect sizes for overall, marginal and regular employment in the main sample.)

Table 2 breaks down the analysis by subgroups. Women from group A were not regularly employed before separation but were to a substantial part marginally employed at t_{-730} (treated: 44%; control: 40%; see Fig. 3). The average divorce effect is higher and women exit marginal employment to a significant degree already before the divorce file was opened. Marginal employment is on average 21 pp lower in t_0 than it would be without divorce. This effect does not fade out and stays rather constant even at the three subsequent measure points in t_{365}, t_{730} and t_{1095} (Fig. 3). At the same time, regular employment increases by 13 pp in t_0 due to divorce and even further to 25 pp in t_{1095} (Fig. 4). (See also Figs. 9, 10, 11 in Appendix for the effect size for all three employment types.)

In contrast, women from group B (with strong labor market participation in regular employment in t_{-2189} to t_{-730}) have no significant employment effects compared to the control group, i.e. the employment rates of divorcees and married women do not differ (Table 2 or Fig. 5). That implies that divorce has neither improved nor worsened the employment status of those divorcees in our observation period. Regarding marginal employment, note that the case numbers in group B are very low for marginal employment, so that I do not discuss nor visualize these results. Likewise, I also skipped the visualization of the effect size.

Table 1 Selected baseline covariates used in logit estimation for the propensity score before and after matching (main sample)

	Raw sample			Matched sample		
	Treated	Control	ND	Treated	Control	ND
At t_0						
Age	38.2	36.65	0.22	38.2	38.05	0.02
Age difference to spouse	3.19	3.06	0.03	3.19	3.16	0.01
Marriage duration	14.49	12.67	0.26	14.49	14.48	0
In t_{-729} to t_{-365}						
Childbirth (one if yes)	0.01	0.07	−0.99	0.01	0.01	−0.01
At t_{-730}						
Number of child age 0–2	0.11	0.29	−0.52	0.11	0.12	−0.03
Number of child age 3–5	0.28	0.4	−0.22	0.28	0.29	−0.02
Number of child age 6–15	0.84	0.83	0.02	0.84	0.84	0
Number of child age 16 and older	0.38	0.25	0.17	0.38	0.37	0.01
In t_{-1094} to t_{-730}						
Days schooling (higher)	4.53	2.14	0.07	4.53	3.81	0.02
Days vocational training	2.88	1.45	0.05	2.88	3.98	−0.04
Days care	8.07	5.88	0.04	8.07	7.82	0.01
Days parental leave	176.9	242.4	−0.37	176.9	178.5	−0.01
Days work disability	2.8	0.77	0.12	2.8	1.53	0.08
Days unemployed	16.89	19.78	−0.05	16.89	16.48	0.01
Days employed (marginal)	76.1	77.21	−0.01	76.1	75.81	0
Days employed (regular)	160.6	112.4	0.28	160.6	158.7	0.01
Daily income (marginal)	2.33	2.31	0	2.33	2.33	0
Daily income (regular)	26.47	19.07	0.24	26.47	26.2	0.01
Yearly income (marginal)	563	579	−0.02	563	560	0
Yearly income (regular)	8260	5744	0.23	8260	8289	0
From first record to t_{-730}						
Lifetime days schooling (higher)	354.6	411.3	−0.08	354.6	351.2	0.01
Lifetime days vocational training	481.2	492.3	−0.02	481.2	479.5	0
Lifetime days tenure	2875	2428	0.19	2875	2846	0.01
Lifetime days tenure (yearly)	156.6	144.4	0.12	156.6	155.9	0.01
Case numbers	413	1437		410	1436	

Mean values for selected covariates. ND is the normalized difference: $\frac{x_t - x_c}{\sqrt{\frac{V_{xt} + V_{xc}}{2}}}$ with $V_{xt} = \frac{1}{(N_t - 1)} * \sum_{i=1}^{N_t} (x_{it} - x_t)^2$ (V_{xc} respectively) (Rosenbaum and Rubin 1985).

Matched sample is based on kernel matching (Epanechnikov) with $h = 0.056$. *Lifetime days tenure* (*yearly*) is calculated by: *Lifetime days tenure*/[*age*-16-(*lifetime days school*/365)-(*lifetime days training*/365)]. T_0 is the day the divorce file was opened. Case numbers for the raw and matched sample deviate because some cases are lost by kernel matching under the chosen bandwidth

Finally, if I compare the logit model with the GBM model (Table 2), then I observe almost identical point estimates and similar signs in all estimations. I treat this as a strong sign that my results are robust to different analytical applications (logit model versus random trees) and weighting schemes (kernel matching versus odd weights), thus, robust to misspecification.

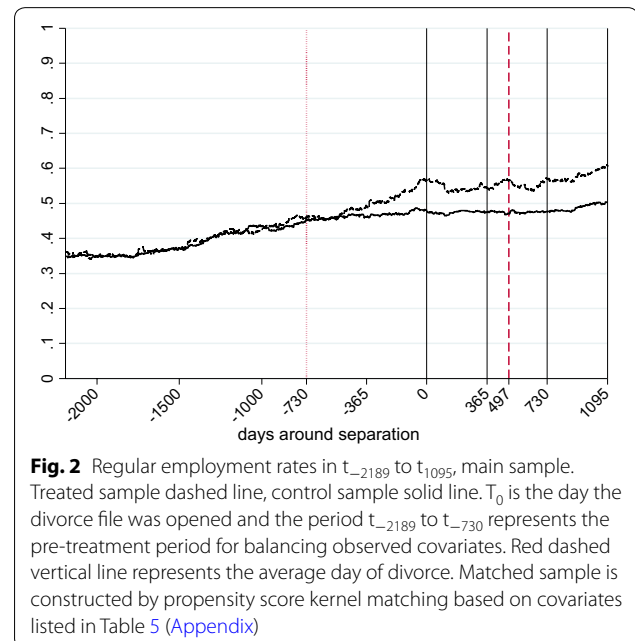## 6.2 Empirical findings—income dynamics with special emphasis on sample selection
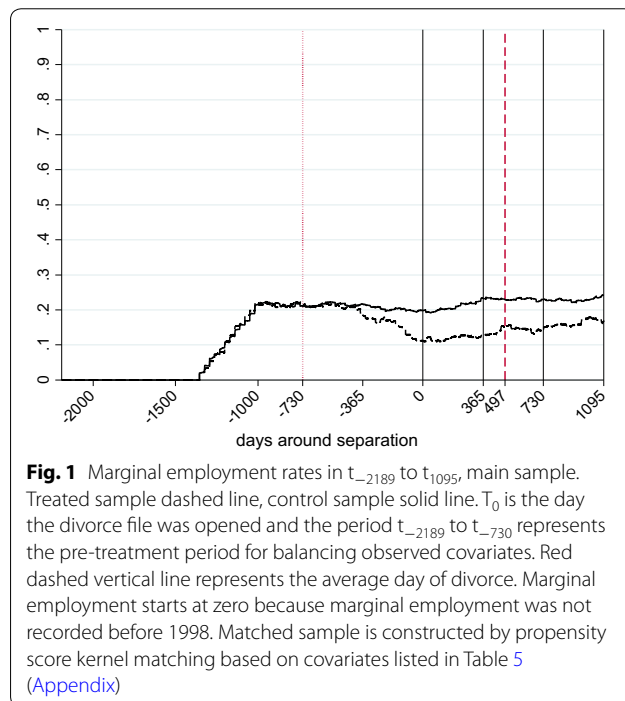
In the presence of sample selection, i.e. non-random selection into employment, the treatment-minus-control difference in incomes might not represent the true causal effect of divorce as long as the non-employed differ systematically in important characteristics from

Table 2 ATT-DiD employment effects in percentage points for main sample, group A and group B at t_0, t_{365}, t_{730} and t_{1095}

	Logit and kernel matching				GBM and odd weighting		
	h bandwidth	Overall employment	Marginal employment	Regular employment	Overall employment	Marginal employment	Regular employment
Main sample							
At day t_0	0.056	− 0.02	− 0.09***	0.08**	− 0.02	− 0.10***	0.09***
At day t_{365}	0.056	− 0.04	− 0.10***	0.06	− 0.03	− 0.10***	0.07*
At day t_{730}	0.056	− 0.01	− 0.09**	0.08*	− 0.01	− 0.09***	0.08**
At day t_{1095}	0.056	0.01	− 0.07*	0.09**	0.00	− 0.08**	0.10**
Group A—women not regularly employed in t_{-2189} to t_{-730}							
At day t_0	0.082	− 0.10	− 0.21***	0.13**	− 0.10*	− 0.22***	0.14***
At day t_{365}	0.082	− 0.08	− 0.25***	0.17***	− 0.06	− 0.24***	0.19***
At day t_{730}	0.082	− 0.02	− 0.23***	0.24***	− 0.01	− 0.23***	0.24***
At day t_{1095}	0.082	0.00	− 0.22***	0.25***	0.00	− 0.22***	0.24***
Group B—women with strong labor market attachment (regular employment) in t_{-2189} to t_{-730}							
At day t_0	0.038	− 0.03	− 0.04	0.00	0.00	− 0.04	0.03
At day t_{365}	0.038	− 0.06	− 0.04	− 0.03	− 0.04	− 0.03	− 0.01
At day t_{730}	0.038	− 0.04	− 0.04	− 0.03	− 0.02	− 0.02	− 0.01
At day t_{1095}	0.038	− 0.04	− 0.01	− 0.04	− 0.01	− 0.02	− 0.01

Cumulative days of regular employment for group A in t_{-2189} to t_{-730} equals zero and for group B is between 967 to 1460 days. Bandwidth h for kernel matching was chosen by leave-one-out cross-validation. Overall employment is marginal and/or regular employment. T_0 is the day the divorce file was opened. Difference-in-difference (DiD) estimation with respect to t_{-730} for all measure points. Matching (and odd weighting) was based on the propensity score. Propensity scores were estimated via a General Boosted Model (GBM) with higher order covariates and interaction terms (depth 3) as well as with a logit model based on covariates in linear form (see Table 5 in Appendix for all applied covariates)

*p < 0.05; **p < 0.01; ***p < 0.001

Fig. 1 Marginal employment rates in t_{-2189} to t_{1095}, main sample. Treated sample dashed line, control sample solid line. T_0 is the day the divorce file was opened and the period t_{-2189} to t_{-730} represents the pre-treatment period for balancing observed covariates. Red dashed vertical line represents the average day of divorce. Marginal employment starts at zero because marginal employment was not recorded before 1998. Matched sample is constructed by propensity score kernel matching based on covariates listed in Table 5 (Appendix)

Fig. 2 Regular employment rates in t_{-2189} to t_{1095}, main sample. Treated sample dashed line, control sample solid line. T_0 is the day the divorce file was opened and the period t_{-2189} to t_{-730} represents the pre-treatment period for balancing observed covariates. Red dashed vertical line represents the average day of divorce. Matched sample is constructed by propensity score kernel matching based on covariates listed in Table 5 (Appendix)
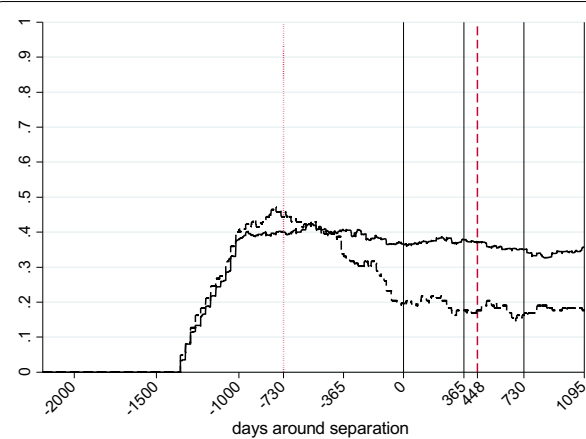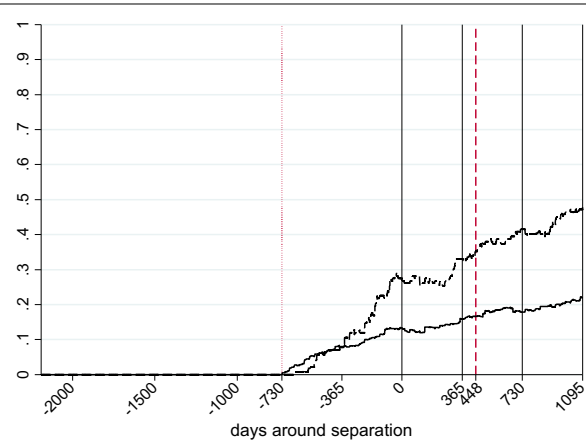
Fig. 3 Marginal employment rates in t_{-2189} to t_{1095}, group A (not regular employed before separation). Treated sample dashed line, control sample solid line. T_0 is the day the divorce file was opened and the period t_{-2189} to t_{-730} represents the pre-treatment period for balancing observed covariates. Red dashed vertical line represents the average day of divorce. Marginal employment starts at zero because marginal employment was not recorded before 1998. Matched sample is constructed by propensity score kernel matching based on covariates listed in Table 5 (Appendix)

Fig. 5 Regular employment rates in t_{-2189} to t_{1095}, group B (strong regular labor force attachment before separation). Treated sample dashed line, control sample solid line. T_0 is the day the divorce file was opened and the period t_{-2189} to t_{-730} represents the pre-treatment period for balancing observed covariates. Red dashed vertical line represents the average day of divorce. Matched sample is constructed by propensity score kernel matching based on covariates listed in Table 5 (Appendix)
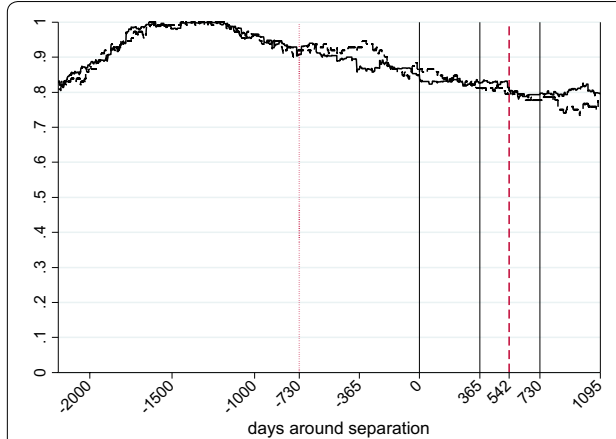
In order to address this issue, I applied a procedure in which the causal treatment effect is not point estimated, but obtained by upper and lower bounds. I derived lower and upper bounds for the set of women who are "always observed", i.e. the share of women who would be employed under the treatment arm and the control arm (Zhang and Rubin 2003; Zhang et al. 2008).[15] Unfortunately, without assumptions, the bounds are usually very large and uninformative and I therefore assumed stochastic dominance, monotonicity and both combined in order to sharpen these bounds.[16]

Fig. 4 Regular employment rates in t_{-2189} to t_{1095}, group A (not regular employed before separation). Treated sample dashed line, control sample solid line. T_0 is the day the divorce file was opened and the period t_{-2189} to t_{-730} represents the pre-treatment period for balancing observed covariates. Red dashed vertical line represents the average day of divorce. Matched sample is constructed by propensity score kernel matching based on covariates listed in Table 5 (Appendix)

---

[15] In the Principal Strata Framework, income is truncated for those who are not employed and women can belong to either one of the following four groups. First, those women who are employed regardless of being treated or not are part of the EE group (always observed, i.e. employed under treatment and control status). Second, women who would be employed when divorced but not employed when married belong to the EN group (employed under treatment and not employed under control status). Third, women who would be non-employed when divorced but would be employed when married belong to the NE group. Lastly, women who would be non-employed whether divorced or not belong to the NN group. The observed employed women (income $Y_i > 0$) from the treatment group consist of the groups EE and EN and the observed employed women from the control group consist of EE and NE. Thus, even controlling for employment is not sufficient since for causal inference treated and control women need to consist of one common set, i.e. only of the EE group. Causal inference is only valid if the EN group from the treated and NE group from the controls are eliminated, such that the income difference is measured at the EE group only, i.e. $\overline{Y}_{EE(treated)} - \overline{Y}_{EE(control)}$ (with $Y_i > 0$). (Zhang and Rubin 2003).

[16] Note, confidence intervals may be constructed to take account of sampling variation according the approach by Imbens and Manski (2004) [for an applied example see Lee (2009)]. I skipped the calculation of confidence intervals since under stochastic dominance all bounds contain "zero" anyway and in those cases where the lower bound was above zero (monotonicity), the plausibility of the assumption is not straightforward.

the employed (Heckman 1979). This is not trivial in my application and Table 5 (columns 3, 4, 5 and 6, Appendix) provides evidence that employed and non-employed women differ sharply. For example, significant predictors of employment are found in childbirth, in the number of toddlers, the education measures, in disability, in parental leave and prior labor market attachment.

Table 3 Sample bounds for the income effect of divorce on daily gross income (regular employment) for the "always observed" under stochastic dominance and/or monotonicity

	$P_{1\|0}$	$P_{1\|1}$	Stochastic dominance		Monotonicity		Stochastic dominance and monotonicity	
			Lower	Upper	Lower	Upper	Lower	Upper
Main sample								
At day t_0	0.479	0.566	-59.25	65.56	-9.17	7.18	0.61	7.18
At day t_{365}	0.474	0.544	-79.22	86.15	-6.68	7.34	1.82	7.34
At day t_{730}	0.477	0.571	-57.85	69.88	-7.60	10.06	2.80	10.06
At day t_{1095}	0.502	0.607	-43.20	53.03	-10.36	8.98	0.89	8.98
Group A—women not regularly employed in t_{-2189} to t_{-730}								
At day t_0	0.134	0.268	-110.45	38.70	-18.41	7.47	-5.29	7.47
At day t_{365}	0.158	0.331	-129.05	37.16	-18.51	7.29	-5.77	7.29
At day t_{730}	0.179	0.415	-96.87	37.78	-17.66	13.88	-1.87	13.88
At day t_{1095}	0.220	0.472	-131.59	48.40	-16.70	15.02	-1.12	15.02
Group B—women with strong labor market attachment (regular employment) in t_{-2189} to t_{-730}								
At day t_0	0.847	0.857	-3.89	11.40	1.82	4.74	3.55	4.74
At day t_{365}	0.827	0.813	-4.04	15.00	5.19	7.54	5.94	7.54
At day t_{730}	0.794	0.777	-3.15	20.77	7.69	10.60	8.72	10.60
At day t_{1095}	0.795	0.768	-5.85	19.76	5.37	10.36	7.34	10.36

Values are in € and $P_{1|0}$ and $P_{1|1}$ are the employment rates (regular employment) for the controls (column 1) and treated (column 2) with $P_{1|0} \equiv \frac{\sum_{i=1}^{N} S_i * (1-T_i) * w_i}{\sum_{i=1}^{N} (1-T_i) * w_i}$ and $P_{1|1} \equiv \frac{\sum_{i=1}^{N} S_i * T_i}{\sum_{i=1}^{N} T_i}$ with $S_i \in (0,1)$ if non-employed or employed and $T_i \in (0,1)$ if control or treated. T_0 is the day the divorce file was opened. Cumulative days of regular employment for group A in t_{-2189} to t_{-730} equals zero and for group B is between 967 to 1460 days. Lower and upper bounds are derived with weights from the logit model and kernel matching. Note that differences in employment rates (between $P_{1|0}$ and $P_{1|1}$) are different to Table 2 (column 4) because in Table 2 I estimated the ATT-DiD. The difference here (between $P_{1|0}$ and $P_{1|1}$) represents the ATT

Stochastic dominance is very likely to hold in the divorce context because it implies that the average daily income of the "always observed" is no less than that of women who are employed under only one treatment arm, i.e. treatment *or* control but not as opposed to treatment *and* control (see footnote 15). To justify that assumption, I assumed that the "always observed" are very likely more motivated, talented or able. As long as these skills transform into higher daily incomes by higher wages and/or more hours worked, this assumption seems reasonable (Zhang et al. 2008; Huber and Mellace 2015).

Table 3 provides the lower and upper bounds for the three groups analyzed and in what follows, I provide a brief example of how they were calculated under the assumption of monotonicity. For monotonicity, the starting point is to calculate the trimming share by using the employment rate for the treated and control group $((P_{1|1} - P_{1|0})/P_{1|1})$. For the main sample at t_0 this results in a trimming value of 15.4% for the employed treated sample, which means that for the upper (lower) bound the lower (upper) part of the (sorted) income distribution is dropped. The income distribution of the employed controls is not trimmed and the average daily gross income is € 55.32 at t_0. For the treated, the average daily gross income is € 62.50 at t_0 for the upper bound (the lower part of the income distribution was dropped) and it is €

46.15 for the lower bound (the upper part of the income distribution was dropped). The bounds under monotonicity are now simply the difference in mean values between the treated and controls.[17]

In Table 3 (column 3 and column 4), I see that under the stochastic dominance assumption the lower and upper bounds contain zero. Hence, I cannot rule out that divorce might only have a pure labor supply effect by encouraging women to enter regular employment while leaving daily earnings unaffected.

For my main sample and group A all bounds are also very large and uninformative. In addition, while for the main sample negative or positive income effects are equally likely, for group A the negative effects are dominating the positive effects (column 3 and 4). Thus, those results highlight that women from group A (with many being mothers, see Table 6, Appendix) are very likely disadvantaged in terms of income effects, when it comes to divorce. One might argue, that this is rooted in the double burden of employment and child rearing because the share of mothers is highest in this sample.

---

[17] Note, the calculation under the assumption of stochastic dominance as well as monotonicity and stochastic dominance combined are different (see Zhang and Rubin 2003; Zhang et al. 2008; Huber and Mellace 2015).

For group B, however, the bounds are narrower with the lower bound quite close to zero. The width of the bounds is reasonably small and is (in comparison to the main sample and group A) suggestive of positive income effects because the negative region of the bound is small compared to the positive region. The evidence provided shows that the actual causal effect on daily income caused by divorce under stochastic dominance is somewhere between € − 3.89 and € 11.40 at t_0 in my sample. Note that the bounds are slightly narrower (but with the lower bound still below zero) if I apply the weights from the GBM model (results are not shown in the table). Figures 12 and 13 (Appendix) provide an overview of lower and upper bounds for group B and for each day in the observation period.

If I also assume monotonicity then I am subsequently able to combine both assumptions, which delivers sharper bounds well above zero for the main sample and group B (Table 3, column 7 and 8). This indicates a causal impact of divorce on individual labor earnings in the samples. However, although such results are promising, the assumption of positive (negative) monotonicity requires that the treatment always leads to higher (lower) labor market participation and rules out increased (decreased) reservation wages (Zhang and Rubin 2003). This assumption might be too strong in the context of divorce and the discussion in the theoretical part has shown that individual labor market exits due to divorce are plausible. Therefore, the plausibility of monotonicity might be too much of a stretch because it rules out the existence of women who would be non-employed when divorced but employed when married.

7 Sensitivity analysis

Until now, I derived the employment effects under the premise that unobserved confounders do not exist or are not relevant. In this section, I scrutinize this assumption and consider selection on unobserved covariates (hidden bias). The reason is that if treated and control units differ in unobserved confounders, i.e. characteristics that simultaneously influence treatment assignment and employment, then the estimated divorce effect is biased.

In Table 4, I display the e^γ values and the respective significance levels for the main sample and group A. I skipped group B because a sensitivity analysis for non-significant employment effects (Table 2, last panel) is not meaningful (Becker and Caliendo 2007).

What is e^γ? The idea of the sensitivity analysis is to check whether the CIA holds. For that reason, I explicitly imply unobserved covariates (hidden bias) and study the influence on the estimated employment effect. Rosenbaum (1995) has shown that the log-odds can be written as a function of observable

Table 4 Sensitivity analysis for unobserved heterogeneity (based on the logit model)

	Marginal employment		Regular employment	
	e^γ	Significance level	e^γ	Significance level
Main sample				
At day t_0	1.70	0.048	2.00	0.046
At day t_{365}	1.65	0.045	1.75	0.048
At day t_{730}	1.52	0.045	1.75	0.045
At day t_{1095}	1.38	0.041	1.80	0.048
Group A—women not regularly employed in t_{-2189} to t_{-730}				
At day t_0	1.38	0.048	1.85	0.048
At day t_{365}	1.60	0.045	2.35	0.049
At day t_{730}	1.68	0.045	2.63	0.046
At day t_{1095}	1.60	0.048	2.50	0.048

T_0 is the day the divorce file was opened. Cumulative days of regular employment for group A in t_{-2189} to t_{-730} equals zero. E^γ and $1/e^\gamma$ provide sharp bounds for the selection into treatment. The hypothetical selection bias (due to unobserved or unmeasured confounders) within the bounds, however, does not drive the confidence intervals of the treatment effects from Table 2 (column 3 and 4) into zero (based on 5% significance level)

characteristics x_i and unobserved characteristics u_i with $F(\beta x_i + \gamma u_i)$. If I denote the treatment (D) probability $P_i = P(D_i = 1|x_i, u_i)$, then the odds ratio for two women i and j are given by:

$$\frac{\frac{P_i}{1-P_i}}{\frac{P_j}{1-P_j}} = \frac{e^{(\beta x_i + \gamma u_i)}}{e^{(\beta x_j + \gamma u_j)}} = e^{(\beta(x_i - x_j) + \gamma(u_i - u_j))}.$$

In the case of a randomized controlled trial, randomization ensures that observed characteristics are $x_i = x_j$ and unobserved characteristics are $u_i = u_j$. Hence, each cancel out so that $e^0 = 1$ remains and both women i and j have the same chance of receiving the treatment (which also implies that no unobserved selection bias exists and the estimated ATT is the true unbiased treatment effect). However, in a study based on administrative data (without being able to randomize women into the control group or treatment group) there is very likely a hidden bias coming from unobserved covariates like marriage quality or the motivation to or not to divorce. In this case, the two women have the same observed characteristics x_i and x_j with $\beta = 0$ (as I can show in Tables 1, 6, 7 and 8) but they very likely differ in unobserved characteristics with $\gamma \neq 0$ and thus, might also differ in the treatment probability. For $\gamma \neq 0$, I can now bound the possible range of the odds ratio by:

$$\frac{1}{e^\gamma} \leq \frac{\frac{P_i}{1-P_i}}{\frac{P_j}{1-P_j}} \leq e^\gamma.$$

With $e^{\gamma} = 1$ the range is simply from 1 to 1 and implies no selection bias but if, for example, $e^{\gamma} = 2$ then the range broadens from ½ to 2 and the odds of the two women could differ up to a factor of 2 or 100%. Intuitively, as the odds ratio differ (and thus, the selection into treatment) the estimated treatment effect and the ATT might be as small as the minimum value (derived for the lower bound) or as high as the maximum value (derived for the upper bound). The task of the sensitivity analysis is to find the point (by slowly increasing γ) where the confidence intervals for the ATT include zero. If e^{γ} close to one already changes the inference about the divorce effect, then the estimates are highly sensitive to hidden bias. However, if the inference is unchanged even for high values of e^{γ}, then the estimated effects are said to be insensitive to hidden bias. This approach does not show that unobserved confounders are present nor that they not exist, but it provides useful information for the discussion to what extent unobserved confounders could alter the treatment effect if they were present (Rosenbaum 1991).

Table 4 highlights that results for regular employment in both samples are relatively insensitive to deviations from the CIA as e^{γ} is ≥ 1.75 which I consider to be large given my observed baseline covariates and the successful randomization (or balance on observed covariates). I can therefore conclude that even large amounts of unobserved heterogeneity would not deteriorate the estimated employment effects in Table 2. Regarding marginal employment, however, the smallest value for e^{γ} is 1.38. Estimated employment effects in Table 2 are therefore much more vulnerable to unobserved covariates that simultaneously influence divorce assignment and labor market participation. Thus, inference about the impact of divorce on marginal employment (at least for t_0 in group A and t_{1095} for the main sample) should be drawn with less confidence.

8 Conclusion

In this paper, I addressed the causal impact of divorce on labor supply and individual income. To that end, I relied on kernel matching and DiD as well as on odd weighting and DiD. I applied two different techniques to estimate the propensity score and can show that the way in which I derived these scores did not affect my estimates. I thus consider my results to be robust to misspecification.

Prior descriptive research had generally shown that divorce leads to an increase in women's employment and individual labor earnings after divorce (Hauser et al. 2016; Bröckel and Andreß 2015). My more causal investigation that differentiates by different types of employment shows a different and more nuanced pattern. First,

I do not find that employment increases after divorce if overall employment is the outcome of interest. However, if overall employment is split into regular and marginal employment, then different employment patterns appear. I find a strong impact of divorce on the type of employment. On average, marginal employment is reduced by approximately 9 pp, while at the same time regular employment increases by 8 pp. The effects are even stronger for women who were not regularly employed in the most recent years preceding separation. For this group, marginal employment is reduced by up to 25 pp while at the same time regular employment soars by 13 pp up to 25 pp in the aftermath of divorce. For women with high labor market attachment a divorce did not affect the employment rate.

Regarding the income estimation, my approach shows that beside a pure labor supply effect a divorce does not seem to have an impact on daily earnings. An exception might be women with a strong labor market attachment because lower bounds for the income effect under stochastic dominance are only slightly negative.[18]

Although I tried my best to adopt a causal approach, remaining caveats must be mentioned. First, I did not know the date when women began to anticipate their divorce and when the "treatment" exactly began. I assumed that women typically anticipated a subsequent divorce, changing their working life accordingly before it occurred and thus set the baseline day at t_{-730}.

Moreover, while the employment effect strongly depends on the CIA (for an unbiased estimation of the causal effect), the income effect relies on additional assumptions. I addressed the CIA explicitly in a sensitivity analysis and found that in particular employment effects for regular employment are insensitive to unobserved confounders. However, employment effects for marginal employment are much more dependent on the CIA. Income effects rely in particular on the stochastic dominance assumption. If monotonicity is also assumed, then I am able to derive lower bounds for the effect of divorce on daily income that are above zero and thus imply a positive treatment effect. While stochastic dominance seems to be plausible, I did not find convincing arguments that monotonicity applies too.

In addition, the causal estimates are based on women with a file opening in 2002. Since labor markets and institutional settings are not static, the estimated effects do not necessarily apply to earlier or later periods. In particular, due to a maintenance reform in 2008 and various reforms to increase the provision of day care for children

---

[18] Only if one is willing to also assume monotonicity, then the lower bounds for daily gross incomes are positive and in the main sample are between € 0.61 to € 2.8 and in group B (high labor market attachment) between € 3.55 to € 8.72. Notable, in group A (low labor market attachment) they are still below zero (€ −5.77 to € −1.12).

since 2005, it is very likely that employment and income effects are more pronounced in more recent years.

Furthermore, as the pension data only include divorces with pension point adjustments, my sample might be selective and does not represent the total population of all divorcees in Germany in 2002. I, therefore, limit my results to the well-defined population of women with pension rights adjustments in the divorce process (which are roughly two thirds of the total divorce population).

Abbreviations

ATT: Average treatment effect on the treated; CIA: Conditional independence assumption; DiD: Difference-in-difference; VA: Versorgungsausgleichsstatistik (dataset); GBM: General boosted models; KS: Kolmogorov–Smirnov-statistic; ND: Normalized difference; pp: Percentage points; SUTVA: Stable unit treatment value assumption; VSKT: Versicherungskontenstichprobe (dataset).

Authors' contributions

Single authorship. The author read and approved the final manuscript.

Appendix

See Tables 5, 6, 7 and 8.

Figures for employment effects and income effects (Figs. 6, 7, 8, 9, 10, 11, 12 and 13).

Table 5 Various logit estimations on the treatment indicator and employment (regular) status on all baseline covariates

Variable	Logit raw sample on treatment indicator	Logit matched sample on treatment indicator	Logit matched sample on regular employment at t_0	Logit matched sample on regular employment at t_{365}	Logit matched sample on regular employment at t_{730}	Logit matched sample on regular employment at t_{1095}
Treatment indicator (divorce)	na	na	0.5733***	0.4268**	0.5186***	0.5546***
Age at t_0	− 0.0016	0.0068	− 0.0417	− 0.0142	0.0194	− 0.0367
Age difference to spouse at t_0	0.0042	0.0010	0.0347	0.0146	0.0283	0.0007
Marriage duration at t_0	0.0315	− 0.0114	0.038	− 0.0001	− 0.0263	0.0209
Child before wedding (one if yes)	0.3437	− 0.0650	0.0059	− 0.1006	− 0.065	− 0.2261
Number of child 0–2 at t_{-730}	− 0.6292**	− 0.0745	− 0.3893	− 0.3698	− 0.2919	− 0.232
Number of child 3–5 at t_{-730}	− 0.0813	− 0.0404	− 0.3759	− 0.7683***	− 0.4616**	− 0.3459*
Number of child 6–15 at t_{-730}	− 0.0451	0.0225	0.0654	− 0.2236	− 0.1841	− 0.265*
Number of child 16 and older at t_{-730}	− 0.2278	0.0371	0.0195	0.1552	0.0013	0.0577
Childbirth in t_{-729} to t_{-365} (one if yes)	− 2.6353***	− 0.0107	− 3.2759***	− 2.7016***	− 1.2672**	0.0189
Days schooling (higher) in t_{-2189} to t_{-1825}	− 0.0005	0.0007	− 0.0005	− 0.0014	0.0002	0
Days schooling (higher) in t_{-1824} to t_{-1460}	0.0023	− 0.0005	− 0.0041	− 0.0024	0.0006	− 0.0024
Days schooling (higher) in t_{-1459} to t_{-1095}	− 0.0006	0.0003	0.0019	− 0.0006	− 0.0008	0.0042

Table 5 (continued)

Variable	Logit raw sample on treatment indicator	Logit matched sample on treatment indicator	Logit matched sample on regular employment at t_0	Logit matched sample on regular employment at t_{365}	Logit matched sample on regular employment at t_{730}	Logit matched sample on regular employment at t_{1095}
Days schooling (higher) in t_{-1094} to t_{-730}	0.0019	0.0004	0.001	0.0018	−0.0005	−0.0016
Days vocational training in t_{-2189} to t_{-1825}	0.0012	−0.0008	−0.0025	−0.0031	−0.0036	−0.002
Days vocational training in t_{-1824} to t_{-1460}	0.0013	0.0021	0.0033	−0.0017	0.004	0.0038
Days vocational training in t_{-1459} to t_{-1095}	−0.0061	−0.0024	−0.003	0.0022	−0.0056	−0.005
Days vocational training in t_{-1094} to t_{-730}	0.0067*	−0.0004	0.0096**	0.0069	0.0111***	0.009**
Days care in t_{-2189} to t_{-1825}	−0.0012	0.0001	0.0063	0.0042	0.0002	−0.0027
Days care in t_{-1824} to t_{-1460}	0.0025	−0.0003	−0.0091	−0.0031	0	0.005
Days care in t_{-1459} to t_{-1095}	−0.0028	0.0011	0.0037	0.0003	0.0024	−0.0008
Days care in t_{-1094} to t_{-730}	0.0022	−0.0006	−0.0021	0.0015	−0.0018	−0.0012
Days parental leave in t_{-2189} to t_{-1825}	−0.0012	−0.0002	0.0037**	0.0035*	0.0022	0.0014
Days parental leave in t_{-1824} to t_{-1460}	0.0013	−0.0002	−0.0034	−0.0013	−0.0001	0.0004
Days parental leave in t_{-1459} to t_{-1095}	−0.0001	0.0003	−0.0015	−0.0021	−0.0023	0.0002
Days parental leave in t_{-1094} to t_{-730}	−0.0013	0.0000	0.0023	0.0025*	0.0025*	0.0003
Days disabled in t_{-2189} to t_{-1825}	0.0019	0.0020	−0.0048	−0.0126	−0.012	−0.0077
Days disabled in t_{-1824} to t_{-1460}	−0.0006	0.0000	−0.021**	−0.0104	−0.0054	−0.0075
Days disabled in t_{-1459} to t_{-1095}	0.0041	−0.0049	0.0063	0.0029	0.0067	−0.001
Days disabled in t_{-1094} to t_{-730}	0.0082*	0.0056	−0.0045	−0.0075	−0.0026	−0.0016
Days unemployed in t_{-2189} to t_{-1825}	0.0011	−0.0002	−0.0009	−0.002	−0.0028*	−0.0015
Days unemployed in t_{-1824} to t_{-1460}	−0.0010	0.0002	0.0013	0.0023	−0.0011	−0.0007
Days unemployed in t_{-1459} to t_{-1095}	0.0014	0.0002	−0.0033*	−0.0027	0.0014	0.001
Days unemployed in t_{-1094} to t_{-730}	−0.0014	−0.0001	0.0015	0.0017	0.0006	0.0009
Days marginal employed in t_{-1459} to t_{-1095}	0.0045	0.0000	0.0018	0.0007	0.0021	−0.0023
Days marginal employed in t_{-1094} to t_{-730}	0.0014	0.0001	0.0054**	0.0037*	0.0034*	0.0057***

Table 5 (continued)

Variable	Logit raw sample on treatment indicator	Logit matched sample on treatment indicator	Logit matched sample on regular employment at t_0	Logit matched sample on regular employment at t_{365}	Logit matched sample on regular employment at t_{730}	Logit matched sample on regular employment at t_{1095}
Days regular employed in t_{-2189} to t_{-1825}	−0.0001	−0.0001	0.0006	0.0004	−0.0002	−0.0011
Days regular employed in t_{-1824} to t_{-1460}	0.0010	−0.0002	0.0013	0.0029	0.0014	0.0003
Days regular employed in t_{-1459} to t_{-1095}	−0.0003	0.0000	0.0019	−0.002	0.0002	0.0029
Days regular employed in t_{-1094} to t_{-730}	0.0013	0.0003	0.0055**	0.0038*	0.0042*	0.0011
Daily marginal income in t_{-1459} to t_{-1095}	0.0238	0.0047	0.0208	0.0102	0.0375	0.0038
Daily marginal income in t_{-1094} to t_{-730}	0.0390	−0.0010	0.0208	0.0032	0.0357	0.016
Daily regular income in t_{-2189} to t_{-1825}	−0.0042	0.0018	0.0039	0.0156*	0.0094	−0.0044
Daily regular income in t_{-1824} to t_{-1460}	0.0037	−0.0002	0.0055	0.008	0.0045	0.0084
Daily regular income in t_{-1459} to t_{-1095}	0.0032	0.0003	0.0141	0.0067	0.0039	−0.0002
Daily regular income in t_{-1094} to t_{-730}	0.0073	0.0026	0.0172**	0.0158*	0.0079	0.0156*
Yearly marginal income in t_{-1459} to t_{-1095}	−0.0006	−0.0001	−0.0001	0	−0.0004	0.0002
Yearly marginal income in t_{-1094} to t_{-730}	−0.0002	0.0000	−0.0007**	−0.0004	−0.0005	−0.0007**
Yearly regular income in t_{-2189} to t_{-1825}	0.0000	0.0000	0	0	0	0
Yearly regular income in t_{-1824} to t_{-1460}	0.0000	0.0000	−0.0001	−0.0001	0	0
Yearly regular income in t_{-1459} to t_{-1095}	0.0000	0.0000	−0.0001*	0	0	−0.0001*
Yearly regular income in t_{-1094} to t_{-730}	0.0000	0.0000	0	0	0	0.0001
Lifetime days schooling (higher)	−0.0001	0.0000	0.0002	0.0003*	0.0001	0.0001
Lifetime days vocational training	0.0000	0.0000	0.0003	0.0003	0.0005**	0.0003
Lifetime days tenure	0.0000	0.0000	0	0.0001	0.0002*	0.0002*
Lifetime days tenure yearly	0.0004	−0.0003	0.003	0.0022	−0.0006	−0.0014

Table 5 (continued)

Variable	Logit raw sample on treatment indicator	Logit matched sample on treatment indicator	Logit matched sample on regular employment at t_0	Logit matched sample on regular employment at t_{365}	Logit matched sample on regular employment at t_{730}	Logit matched sample on regular employment at t_{1095}
_cons	− 1.3570	− 0.1059	− 1.5716	− 1.7672*	− 2.4289**	− 0.3857
N	1850	1846	1846	1846	1846	1846
Pseudo R^2	0.0792	0.0033	0.3482	0.2904	0.2351	0.1928

Training is occupational training. School is attending higher school. Following Sianesi (2004) the Pseudo R^2 in column 2 signals that the distribution of covariates is very well balanced within the matched sample because the value is very low. Thus, this is another sign that matching was successful (for comparison see value in column 1 for the raw sample). *Lifetime school*, *Lifetime training* and *Lifetime tenure* are cumulated days since age 15 to $t_{−730}$. *Lifetime tenure yearly* is calculated by: *Lifetime tenure*/[*age*-16-(*Lifetime school*/365) − (*Lifetime training*/365)]

*p < 0.05; **p < 0.01; ***p < 0.001

Table 6 Selected baseline covariates used in logit estimation for the propensity score before and after matching (group A)

	Raw sample			Matched sample		
	Treated	Control	ND	Treated	Control	ND
At t_0						
Age	38.33	36.73	0.25	38.33	38.06	0.04
Age difference to spouse	3.33	3.19	0.03	3.33	3.23	0.02
Marriage duration	14.86	13.00	0.27	14.86	14.62	0.03
In $t_{−729}$ to $t_{−365}$						
Childbirth (one if yes)	0.01	0.05	− 0.51	0.01	0.01	− 0.03
At $t_{−730}$						
Number of child 0–2	0.14	0.28	− 0.37	0.14	0.16	− 0.04
Number of child 3–5	0.37	0.50	− 0.23	0.37	0.39	− 0.04
Number of child 6–15	1.06	1.08	− 0.02	1.06	1.08	− 0.02
Number of child 16 and older	0.37	0.22	0.21	0.37	0.35	0.03
In $t_{−1094}$ to $t_{−730}$						
Days schooling (higher)	6.50	3.78	0.06	6.50	4.93	0.04
Days vocational training	0.54	1.14	− 0.10	0.54	0.79	− 0.04
Days care	11.70	9.89	0.03	11.70	13.75	− 0.03
Days parental leave	223.29	261.79	− 0.22	223.29	229.19	− 0.03
Days work disability	0.00	0.00	NA	0.00	0.00	NA
Days unemployed	6.78	12.36	− 0.13	6.78	7.64	− 0.02
Days employed (marginal)	150.00	113.09	0.23	150.00	136.66	0.08
Days employed (regular)	0.00	0.00	NA	0.00	0.00	NA
Daily income (marginal)	3.76	3.14	0.16	3.76	3.49	0.07
Daily income (regular)	0.00	0.00	NA	0.00	0.00	NA
Yearly income (marginal)	1114.17	870.26	0.18	1114.17	1018.44	0.07
Yearly income (regular)	0.00	0.00	NA	0.00	0.00	NA
From first record to $t_{−730}$						
Lifetime days schooling (higher)	309.07	419.36	− 0.19	309.07	327.55	− 0.03
Lifetime days vocational training	425.54	401.22	0.05	425.54	412.22	0.03
Lifetime days tenure	1517.30	1406.73	0.08	1517.30	1457.04	0.04
Lifetime days tenure (yearly)	79.69	78.78	0.01	79.69	77.26	0.04
Case numbers	144	654		142	642	

Mean values for selected covariates (baseline). ND is the normalized difference: $\frac{x_t - x_c}{\sqrt{\frac{V_{xt} + V_{xc}}{2}}}$ with $V_{xt} = \frac{1}{(N_t - 1)} * \sum_{i=1}^{N_t} (x_{it} - x_t)^2$ (V_{xc} respectively) (Rosenbaum and Rubin 1985). Matched sample is based on kernel matching (Epanechnikov) with $h = 0.082$. *Lifetime days tenure* (*yearly*) is calculated by: *Lifetime days tenure*/[*age*-16-(*lifetime days school*/365)-(*lifetime days training*/365)]. T_0 is the day the divorce file was opened. Case numbers for the raw and matched sample deviate because some cases are lost by kernel matching under the chosen bandwidth

Table 7 Selected baseline covariates used in logit estimation for the propensity score before and after matching (group B)

	Raw sample			Matched sample		
	Treated	**Control**	**ND**	**Treated**	**Control**	**ND**
At t_0						
Age	40.41	39.89	0.07	40.41	40.36	0.01
Age difference to spouse	2.73	2.67	0.02	2.73	2.84	− 0.02
Marriage duration	16.50	15.52	0.13	16.50	16.53	0.00
In t_{-729} to t_{-365}						
Childbirth (one if yes)	0.00	0.00	NA	0.00	0.00	NA
At t_{-730}						
Number of child 0–2	0.04	0.11	− 0.39	0.04	0.03	0.04
Number of child 3–5	0.04	0.09	− 0.28	0.04	0.03	0.04
Number of child 6–15	0.66	0.67	− 0.01	0.66	0.62	0.04
Number of child 16 and older	0.51	0.52	− 0.01	0.51	0.53	− 0.03
In t_{-1094} to t_{-730}						
Days schooling (higher)	0.00	0.00	NA	0.00	0.00	NA
Days vocational training	0.00	0.00	NA	0.00	0.00	NA
Days care	1.86	1.20	0.03	1.86	2.78	− 0.05
Days parental leave	88.39	133.00	− 0.29	88.39	82.04	0.04
Days work disability	3.28	1.91	0.09	3.28	3.21	0.01
Days unemployed	6.74	8.69	− 0.05	6.74	5.18	0.04
Days employed (marginal)	3.12	4.17	− 0.04	3.12	1.62	0.05
Days employed (regular)	344.89	333.45	0.16	344.89	346.30	-0.02
Daily income (marginal)	0.14	0.21	− 0.06	0.14	0.08	0.06
Daily income (regular)	56.00	54.83	0.04	56.00	56.58	− 0.02
Yearly income (marginal)	30.13	24.73	0.02	30.13	15.45	0.05
Yearly income (regular)	19,681.67	18,539.90	0.11	19,681.67	19,881.25	− 0.02
From first record to t_{-730}						
Lifetime days schooling (higher)	393.00	348.57	0.06	393.00	398.00	− 0.01
Lifetime days vocational training	534.91	585.07	− 0.10	534.91	529.90	0.01
Lifetime days tenure	4931.96	4989.86	− 0.02	4931.96	4960.00	− 0.01
Lifetime days tenure (yearly)	250.76	261.38	− 0.13	250.76	251.46	− 0.01
Case numbers	134	328		112	293	

Mean values for selected covariates (baseline). ND is the normalized difference: $\frac{x_t - x_c}{\sqrt{\frac{V_{xt} + V_{xc}}{2}}}$ with $V_{xt} = \frac{1}{(N_t - 1)} * \sum_{i=1}^{N_t} (x_{it} - \bar{x}_t)^2$ (V_{xc} respectively) (Rosenbaum and Rubin 1985). Matched sample is based on kernel matching (Epanechnikov) with $h = 0.038$. *Lifetime days tenure* (*yearly*) is calculated by: *Lifetime days tenure*/[*age*-16-(*lifetime days school*/365)-(*lifetime days training*/365)]. T_0 is the day the divorce file was opened. Case numbers for the raw and matched sample deviate because some cases are lost by kernel matching under the chosen bandwidth

Table 8 Balance quality for raw, matched and weighted sample

	Number		Normalized difference		Kolmogorov–Smirnov	
	Treated	Control	Max ND	Mean ND	Max KS	Mean KS
Main sample						
Raw sample	410	1436	0.988	0.125	0.186	0.061
Logit model and kernel matching	410	956	0.078	0.012	0.066	0.020
GBM model and odd weighting	410	1104	0.229	0.037	0.051	0.026
Group A—women not regularly employed in t_{-2189} to t_{-730}						
Raw sample	142	642	0.510	0.135	0.163	0.044
Logit model and kernel matching	142	430	0.083	0.036	0.094	0.023
GBM model and odd weighting	142	428	0.203	0.065	0.088	0.024
Group B—women with strong labor market attachment (regular employment) in t_{-2189} to t_{-730}						
Raw sample	112	293	0.691	0.111	0.151	0.041
Logit model and kernel matching	112	177	0.079	0.031	0.113	0.033
GBM model and odd weighting	112	160	0.154	0.045	0.099	0.029

Numbers in column 1 and 2 refer to the effective sample size. Numbers for the control group are based on either the kernel weights or the odd weights. Briefly, the effective sample size gives an estimate of the number of controls that are comparable to the treatment group. For an introduction, see for example Ridgeway et al. (2015: 9)
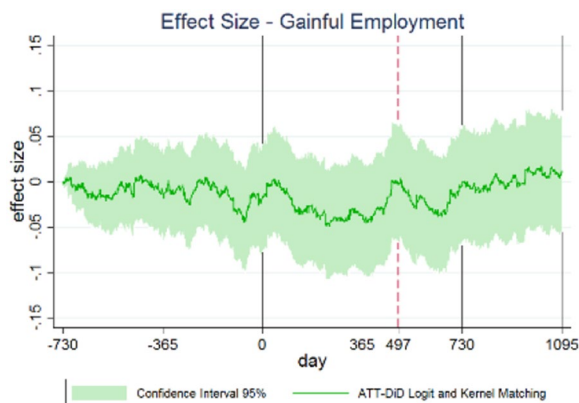
Fig. 6 Effect size for overall employment, main sample. T_0 is the day the divorce file was opened. Red dashed vertical line represents the average day of divorce. The shaded area represents the 95% confidence interval
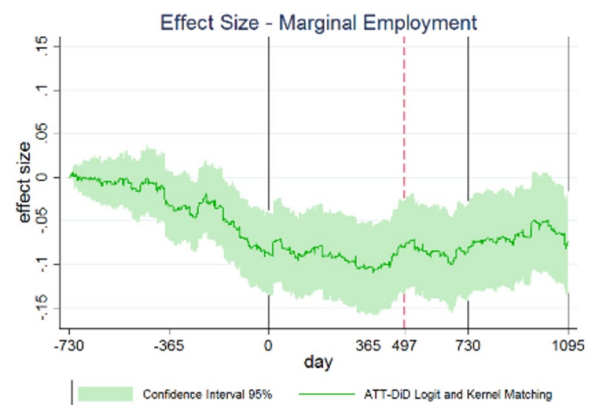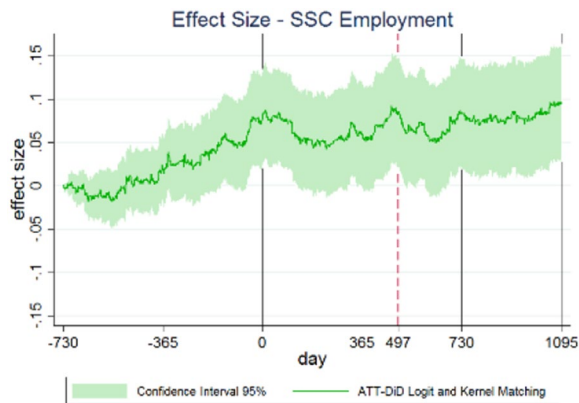

Fig. 7 Effect size for marginal employment, main sample. T_0 is the day the divorce file was opened. Red dashed vertical line represents the average day of divorce. The shaded area represents the 95% confidence interval

Fig. 8 Effect size for regular employment, main sample. T_0 is the day the divorce file was opened. Red dashed vertical line represents the average day of divorce. SSC means employment with social security contribution, i.e. regular employment. The shaded area represents the 95% confidence interval
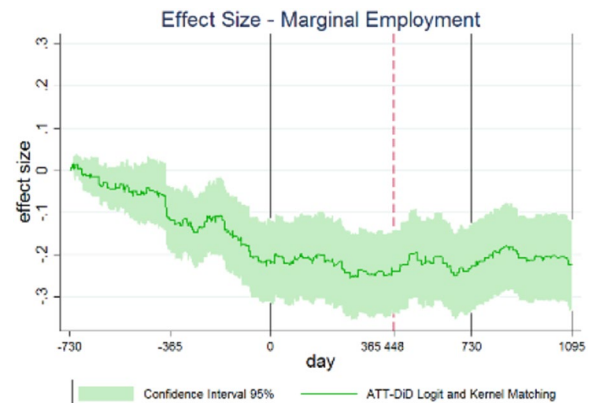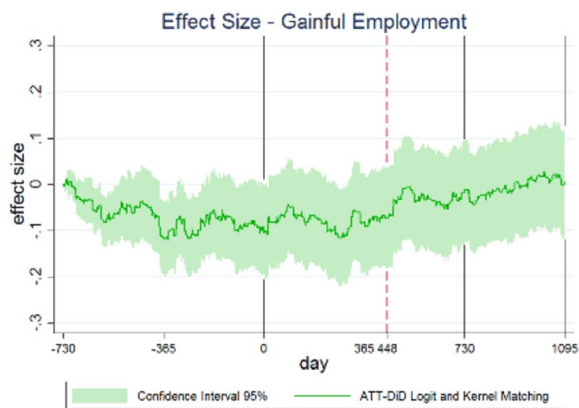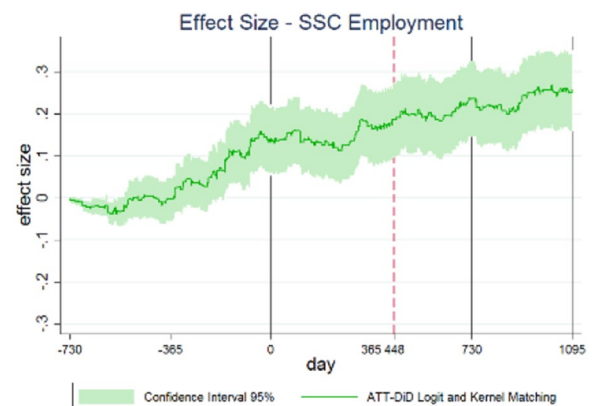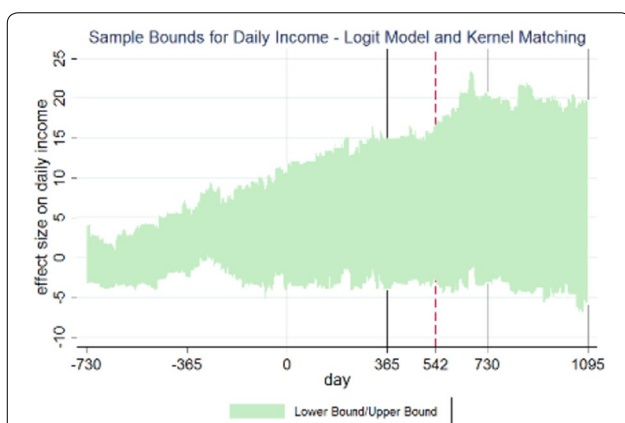

Fig. 10 Effect size for marginal employment, group A. T_0 is the day the divorce file was opened. Red dashed vertical line represents the average day of divorce. The shaded area represents the 95% confidence interval

Fig. 9 Effect size for overall employment, group A. T_0 is the day the divorce file was opened. Red dashed vertical line represents the average day of divorce. The shaded area represents the 95% confidence interval

Fig. 11 Effect size for regular employment, group A. T_0 is the day the divorce file was opened. Red dashed vertical line represents the average day of divorce. SSC means employment with social security contribution, i.e. regular employment. The shaded area represents the 95% confidence interval
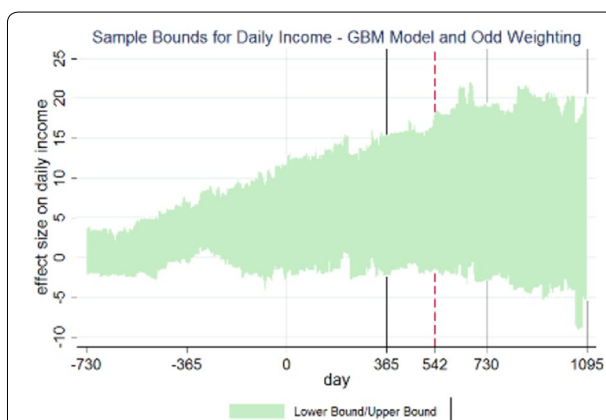
Fig. 12 Lower and upper bound for daily income based on weights from the logit model and derived under stochastic dominance, group B. T_0 is the day the divorce file was opened. Red dashed vertical line represents the average days of divorce. Lower and upper bounds of daily incomes for the subgroup of women with strong labor market attachment in regular employment while married were derived by the principal strata framework (Zhang and Rubin 2003) under the assumption of stochastic dominance

Fig. 13 Lower and upper bound for daily income based on weights from the GBM model and derived under stochastic dominance, group B. T_0 is the day the divorce file was opened. Red dashed vertical line represents the average days of divorce. Lower and upper bounds of daily incomes for the subgroup of women with strong labor market attachment in regular employment while married were derived by the principal strata framework (Zhang and Rubin 2003) under the assumption of stochastic dominance

References

Becker, S., Caliendo, M.: Sensitivity analysis for average treatment effects. Stata J. **7**(1), 71–83 (2007). https://doi.org/10.1177/1536867X0700700104

Becker, G.: A treatise on the family. National Bureau of Economic Research, Inc (1981). https://EconPapers.repec.org/RePEc:nbr:nberbk:beck81-1

Bröckel, M., Andreß, H.J.: The economic consequences of divorce in Germany: what has changed since the turn of the millennium? Comp. Popul. Stud. **40**(3), 277–312 (2015). https://doi.org/10.12765/CPoS-2015-04en

Burkhauser, R.V., Duncan, G.J., Hauser, R., Berntsen, R.: Wife or Frau, women do worse: a comparison of men and women in the United States and Germany after marital dissolution. Demography **28**(3), 353–360 (1991). https://doi.org/10.2307/2061461

Caliendo, M., Künn, S., Uhlendorff, A.: Marginal employment, unemployment duration and job match quality. IZA Discussion Papers 6499, Institute for the Study of Labor (2012)

DiPrete, T.A., McManus, P.A.: Family change, employment transitions, and the welfare state: Household income dynamics in the United States and Germany. Am. Sociol. Rev. **65**(3), 343–370 (2000)

Engstler, H., Menning, S.: Families in Germany—facts and figures. Federal Ministry for Family Affairs, Senior Citizens, Women and Youth, Berlin (2004)

Eurostat: Divorce indicators demo_ndivind (2018). http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_ndivind&lang=en. Accessed 4 Aug 2018

Gangl, M., Ziefle, A.: Motherhood, labor force behavior, and women's careers: an empirical assessment of the wage penalty for motherhood in Britain, Germany, and the United States. Demography **46**(2), 341–369 (2009). https://doi.org/10.1353/dem.0.0056

Hauser, R., Burkhauser, R.V., Couch, K.A., Bayaz-Ozturk, G.: Wife or Frau, women still do worse: a comparison of men and women in the United States and Germany after union dissolutions in the 1990s and 2000s. Working Paper 2016-39, University of Connecticut, Department of Economics (2016)

Heckman, J.J.: Sample selection bias as a specification error. Econometrica **47**(1), 153 (1979). https://doi.org/10.2307/1912352

Huber, M., Mellace, G.: Sharp bounds on causal effects under sample selection. Oxf. Bull. Econ. Stat. **77**(1), 129–151 (2015). https://doi.org/10.1111/obes.12056

Hummelsheim, D.: Germany: will the male breadwinner model survive? In: Hummelsheim, D., Andress, H.-J. (eds.) When marriage ends: economic and social consequences of partnership dissolution, pp. 51–77. Edward Elgar, Cheltenham (2009)

Imbens, G.W., Manski, C.F.: Confidence intervals for partially identified parameters. Econometrica **72**(6), 1845–1857 (2004). https://doi.org/10.1111/j.1468-0262.2004.00555.x

Jenkins, S.P.: Marital splits and income changes over the longer term. ISER working paper series 2008–07, Institute for Social and Economic Research (2008)

Keck, W., Radenacker, A., Brüggmann, D., Kreyenfeld, M., Mika, T.: Statutory pension insurance accounts and divorce: a new scientific use file. Jahrbücher für Nationalökonomie und Statistik (JBNST) (2019). https://doi.org/10.1515/jbnst-2019-0064

Killingsworth, M.R., Heckman, J.J.: Female labor supply: a survey. In: Ashenfelter, O., Layard, R. (eds.) Handbook of labor economics, vol. 1, pp. 103–204. Elsevier, Amsterdam (1987)

Kruse, E.: Empirische Ergebnisse zur Erwerbsbeteiligung älterer Versicherter anhand der Statistiken der gesetzlichen Rentenversicherung DRV (11-12/2007), 716–749 (2007)

Lechner, M.: The estimation of causal effects by difference-in-difference methods. Found. Trends Econom. **4**(3), 165–224 (2011). https://doi.org/10.1561/0800000014

Lee, D.S.: Training, wages, and sample selection: estimating sharp bounds on treatment effects. Rev. Econ. Stud. **76**(3), 1071–1102 (2009). https://doi.org/10.1111/j.1467-937x.2009.00536.x

Mantel, N., Haenszel, W.: Statistical aspects of the analysis of data from retrospective studies of disease. J. Natl Cancer Inst. (1959). https://doi.org/10.1093/jnci/22.4.719

McCaffrey, D.F., Griffin, B.A., Almirall, D., Slaughter, M.E., Ramchand, R., Burgette, L.F.: A tutorial on propensity score estimation for multiple treatments using generalized boosted models. Stat. Med. **32**(19), 3388–3414 (2013). https://doi.org/10.1002/sim.5753

McKeever, M., Wolfinger, N.H.: Reexamining the economic costs of marital disruption for women. Soc. Sci. Q. **82**(1), 202–217 (2001). https://doi.org/10.1111/0038-4941.00018

Mueller, R.E.: The effect of marital dissolution on the labour supply of males and females: evidence from Canada. J. Socio-Econ. **34**, 787–809 (2005)

Özcan, B., Breen, R.: Marital instability and female labor supply. Ann. Rev. Sociol. **38**(1), 463–481 (2012). https://doi.org/10.1146/annurev-soc-071811-145457

Raz-Yurovich, L.: Divorce penalty or divorce premium? A longitudinal analysis of the consequences of divorce for men's and women's economic activity. Eur. Sociol. Rev. **29**(2), 373–385 (2011). https://doi.org/10.1093/esr/jcr073

Ridgeway, G., McCaffrey, D., Morral, A., Burgette, L., Griffin, A.: Toolkit for weighting and analysis of nonequivalent groups: a tutorial for the R TWANG package Santa Monica, CA, RAND Corporation (2015). https://www.rand.org/pubs/tools/TL136z1.html

Rosenbaum, P.R.: Discussing hidden bias in observational studies. Ann. Intern. Med. **115**(11), 901–905 (1991). https://doi.org/10.7326/0003-4819-115-11-901

Rosenbaum, P.R.: Observational studies. Springer series in statistics. Springer, Berlin (1995). ISBN 978-1-4757-3692-2

Rosenbaum, P.R., Rubin, D.B.: Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. Am. Stat. **39**(1), 33 (1985). https://doi.org/10.2307/2683903

Seifert, H.: Non-regular employment in Germany. In: Japan Institute for Labour Policy and Training (ed.) Non-regular employment issues and challenges common to the major developed countries, pp. 69–85. The Japan Institute for Labour Policy and Training, Tokyo (2011)

Sianesi, B.: An evaluation of the Swedish system of active labor market programs in the 1990s. Rev. Econ. Stat. **86**(1), 133–155 (2004). https://doi.org/10.1162/003465304323023723

Spieß, C.K.: Vereinbarkeit von Familie und Beruf – wie wirksam sind deutsche Care Policies"? Perspektiven der Wirtschaftspolitik **12**(Supplement), 871 (2011). https://doi.org/10.1111/j.1468-2516.2011.00354.x

Stegmann, M., Himmelreicher, R.K.: Die Versicherungskontenstichprobe als scientific use file DRV-Schriften (Band 79) (2008)

Steiner, V., Wrohlich, K.: Household taxation, income splitting and labor supply incentives—a microsimulation study for Germany. CESifo Econ. Stud. **50**(3), 541–568 (2004). https://doi.org/10.1093/cesifo/50.3.541

Stuart, E.A.: Matching methods for causal inference: a review and a look forward. Stat. Sci. **25**(1), 1–21 (2010). https://doi.org/10.1214/09-STS313

Tach, L.M., Eads, A.: Trends in the economic consequences of marital and cohabitation dissolution in the United States. Demography **52**(2), 401–432 (2015). https://doi.org/10.1007/s13524-015-0374-5

Tamborini, C.R., Couch, K.A., Reznik, G.L.: Long-term impact of divorce on women's earnings across multiple divorce windows: a life course perspective. Adv. Life Course Res. **26**, 44–59 (2015). https://doi.org/10.1016/j.alcr.2015.06.001

Uunk, W.: The economic consequences of divorce for women in the European Union: the impact of welfare state arrangements. Eur. J. Popul. **20**(3), 251–285 (2004). https://doi.org/10.1007/s10680-004-1694-0

van Damme, M., Kalmijn, M., Uunk, W.: The employment of separated women in Europe: individual and institutional determinants. Eur. Sociol. Rev. **25**(2), 183–197 (2009)

Wrohlich, K., Müller, K.U.: Two steps forward - one step back? Evaluating recent childcare policies in Germany. Beiträge zur Jahrestagung des Vereins für Sozialpolitik. Evidenzbasierte Wirtschaftspolitik - Session: Child Care, No. C02-V02 (2014)

Zhang, J.L., Rubin, D.B.: Estimation of causal effects via principal stratification when some outcomes are truncated by "death". J. Educ. Behav. Stat. **28**(4), 353–368 (2003). https://doi.org/10.3102/10769986028004353

Zhang, J.L., Rubin, D.B., Mealli, F.: Evaluating the effects of job training programs on wages through principal stratification. In: Fomby, T., Carter Hill, R., Millimet, D.L., Smith, J.A., Vytlacil, E.J. (Ed.) Modeling and evaluating treatment effects in econometrics (Advances in Econometrics, Vol. 21), pp. 117–145. Emerald Group Publishing Limited, Bingley (2008). https://doi.org/10.1016/s0731-9053(07)00005-9

Publisher's Note