The Sick Pay Trap

Elisabeth Fevang, *Ragnar Frisch Centre for Economic Research*

Simen Markussen, *Ragnar Frisch Centre for Economic Research*

Knut Røed, *Ragnar Frisch Centre for Economic Research*

In most countries, employers are financially responsible for sick pay during an initial period of a worker's absence spell, after which the public insurance system covers the bill. Based on an empirical evaluation of a quasi-natural experiment in Norway, where pay liability was removed for pregnancy-related absences, we show that the system of short-term pay liability creates a sick pay trap: firms are discouraged from letting long-term sick workers back into work since they then face the financial risk associated with subsequent relapses. We present evidence indicating that this disincentive effect is both statistically and economically significant.

I. Introduction

Based on extensive reviews of disability prevention experiences in 13 countries, the Organization for Economic Cooperation and Development (OECD 2010, 125) argues that "employers are key players in preventing health problems at work and facilitating a swift return to work for people absent from work due to sickness." But, while there is ample em-

pirical evidence regarding the responsiveness of absenteeism with respect to worker incentives (Henreksson and Persson 2004; Johansson and Palme 2005; Ziebarth 2009; Ziebarth and Karlsson 2009; D'Amuri 2011), there is little evidence regarding the impact of firm incentives. OECD (2010, 133) notes that countries where employers are responsible for a large share of their employees' sick pay costs tend to have much lower absence rates than countries where employers can pass the costs on to the public purse and also that absenteeism has dropped significantly in the Netherlands and the United Kingdom after a shift of financial responsibility toward employers. Yet, to our knowledge, scientific evidence establishing a causal relationship between firm incentives and worker absenteeism is nonexistent. The design of firm incentives with respect to sick leave prevention also involves a potential trade-off between sick leave and labor-market exclusion: while more extensive pay liability (or experience rating) improves incentives for absence prevention, it may at the same time undermine incentives for employing persons perceived to have a high risk of absence in the first place.

Although the extent of employer co-payment differs sharply across different countries, the incentive structure faced by employers in most industrialized economies typically implies that the firm is responsible for sick pay expenditures during an initial stage of a workers' sick leave but that the national insurance scheme (or another insurer) covers the costs accruing after some duration threshold (OECD 2010, table 5.1).[1] This means that firms do have strong financial incentives to prevent short-term absences. However, in cases where absence spells stretch beyond the co-payment period, employers may (rightly) think that it is not in their interest to facilitate a quick return to work, since the return to work also entails the risk of new short-term absences for which the employers are again financially responsible. Hence, current incentive structures may have the unintended side effect of discouraging employers from exerting appropriate effort to curb long-term absenteeism.

In the present article, we examine empirically the impacts of employers' pay liability by exploiting a reform in the Norwegian sick leave insurance scheme, whereby pay liability for pregnant workers' short-term sick leaves was removed for absences caused by pregnancy-related illnesses (in the

---

[1] While most European countries have public insurance programs for short-term sick leave, typically with replacement rates ranging from 50% to 100% and maximum durations exceeding 1 year (Petterson-Lidbom and Thoursie 2013), there are no federal legal requirements for paid sick leave in the United States. For companies subject to the Family and Medical Leave Act (FMLA), i.e., firms with at least 50 employees, the act does require unpaid sick leave, however. And many US employers offer sick leave pay on a voluntary basis. Also, some states, including California, New Jersey, and New York, have public programs that partially protect workers against the loss of income due to nonoccupational disability.

period prior to transition to maternity leave benefits). The motivation for this reform was that it was feared that the elevated risk of sickness absence associated with pregnancies made employers reluctant toward hiring young women. Markussen et al. (2011) show that the increased risk of absenteeism associated with pregnancies is indeed substantial; the hazard rate of entering into a sick leave spell with a diagnosis predicting full exploitation of the employer's pay liability—often related to back pain, nausea, or anxiety—is raised by a factor of 5 at the onset of a pregnancy and further to a factor of 15 during the last 2–3 months before delivery. The reform thus clearly removed a potentially important disincentive with respect to hiring young female labor, but at the same time it also changed employers' incentives to prevent sick leave among pregnant workers, for example, by reducing productivity expectations/requirements, modifying regular duties, encouraging colleagues to help out with physically demanding tasks, allowing more flexible and less strenuous hours, or investing in technical (strain-reducing) equipment.

On the one hand, the reform made short-term absence—absence spells with durations up to 16 days—less costly for the firms. On the other hand, it also made it less risky to let long-term absent pregnant workers return to work, since the firms no longer were responsible for the sick pay costs associated with subsequent relapses. Hence, the reform offers a neat setting for identifying the impacts of firm incentives. Based on a combination of regression discontinuity (RD) and difference-in-difference (DiD) methodologies, we show that the reform had significant impacts on the affected employees' absence behavior. Using a bivariate hazard rate model (with nonparametric modeling of unobserved heterogeneity), we find that short-term absenteeism originally covered by firms' pay liability rose significantly, while the duration of noncovered long-term spells declined. For example, we estimate that the reform led to a 10% rise in the entry rate into absenteeism but also in a 12% rise in the rate of work resumption at durations exceeding 16 days. Extrapolating our findings to the economy as a whole, we estimate that the 16-day pay liability system reduces the overall number of covered absence days by 11%, while actually raising the number of noncovered days (days exceeding 16 days duration) by 1.6%, compared to a system with no pay liability at all. Our results are exactly as one would expect on the basis of simple economic theory, provided that firms do have some influence on their employees' sick leave behavior. Our findings thus indicate that policy makers indeed may have good reasons to focus on improving employer incentives in their efforts to curb absenteeism.

We also find some evidence indicating that the removal of pay liability for pregnant workers made the school-to-work transition easier for young women. According to our estimates, the reform raised the employment propensity 1 year after graduation by around 1.5 percentage points for

young women in general and by 3.0 percentage points for those who were pregnant at the time of graduation. This suggests that policy makers may have to trade off incentives for sick leave prevention against incentives for employing workers with high expected absenteeism.

II. Related Literature

When an insurance scheme is troubled by moral hazard problems, efficiency considerations suggest that coverage should decrease with duration. This is studied extensively in relation to the design of optimal unemployment insurance, starting from the seminal paper by Shavell and Weiss (1979). The argument is simple. In the presence of moral hazard, there is an inevitable trade-off between insurance and incentives. By reshuffling the benefit schedule to provide lower payments tomorrow and higher payments today, such that expected utility remains constant, agents are given stronger incentives to search for jobs.

Many countries have adopted declining benefit schedules for the unemployed, most often in the form of a single drop after some time in addition to an overall duration limitation (Cahuc and Zylberberg 2004, 143). Maximum duration limitations are also typically in place for sickness insurance payments. However, there are also historical examples of sickness benefits that increase with duration. Johansson and Palme (2005) study a reform in Sweden, where a time-constant replacement rate of 90% was replaced by a time-increasing payment schedule with 65% replacement rate the first 3 days, 80% the next 77 days, and then 90% from day 80. Johansson and Palme (2005) found that the reform changed workers' behavior exactly as the altered incentives would imply. As the cost of short-term absence increased, short-term absenteeism dropped. For long-term absences, however, the return-to-work hazard declined as the risk of relapse raised the expected cost of returning to work. This example illustrates a potential benefit trap: having reached the highest level of replacement, it is not particularly tempting to risk a return to the bottom of the replacement ladder.

In the present article, we focus on the employers' incentives rather than those of the employees. But, provided that employers influence their employees' sick leave behavior, the story is basically the same. If the employer is financially responsible for short-term absence only, the firm obviously has incentives to prefer a single long absence spell over many short ones. And when employees have been on sick leave long enough to have exhausted the firms' pay liability, the financial risk associated with possible relapses may convince the employer not to accommodate a quick return to work.

While it has long been recognized in the literature that incomplete experience rating in unemployment insurance systems incites firms to lay off too many workers (e.g., Topel 1983), we have not been able to find any

empirical evidence regarding the causal relationship between firms' sick leave insurance costs and their workers' absenteeism. One of the closest pieces of evidence we have found is Burkhauser, Schmeiser, and Weathers (2012), which shows that disabled workers in the United States whose work limitations were caused by an accident on the job, and who were, hence, more likely to be covered by the experience-rated workers' compensation program than workers whose injuries were not work related, also were more likely to be offered accommodation by their employer.

Our article also relates to a literature on the labor-demand effects of mandated employer provision of employee benefits in general (see Burkhauser and Daly [2011] for a recent discussion). A contribution to this literature with particular resemblance to our own is Gruber (1994), which found that mandates that raised the costs of insuring female employees of childbearing age in the United States (by including childbirth in health insurance plans) did not adversely affect young women's employment rates, as wages adjusted to compensate for the value of the mandated benefits. It is conceivable, however, that this finding could be reversed in labor markets with less individual wage flexibility—like the Norwegian one.

III. Institutions and Mechanisms

All Norwegian workers are fully insured against sickness absence for up to 1 year, with a 100% replacement ratio.[2] Norway also has a high level of absenteeism. On a typical working day, around 7% of all workers are absent due to sickness. This places Norway among the countries with the world's highest sickness absence rates (see, e.g., Bonato and Lusinyan 2007; Edwards and Greasley 2010). Standard regulations imply that absence spells exceeding 3 days (including weekends, holidays, etc.) need to be certified by a physician. Certification is not formally required until the ninth day for employees in firms participating in a tripartite "inclusive workplace agreement" (IWA) between employers, employees, and the state, although it is common to adhere to the 3-day rule even in these firms.[3] Approximately half of the employees are covered by IWA. In total, around 90% of all absence days in Norway are certified by a physician. The sick pay costs are shared between the employer and the tax payers. The general rule is that

---

[2] There is a ceiling on annual earnings (in 2012) of NOK 500,000 (≈US$85,000). Eligibility requires that the employee have been at work for 4 weeks.

[3] Based on an in-depth study of a large Norwegian IWA-company, Bergsvik, Markussen, and Raaum (2010) show that more than 90% of all self-reported absence spells are 1–3 days long (less than 1% are longer than 5 days), despite the possibility of self-reporting up to 8 days. Statistics reported by the Norwegian Welfare Administration also indicate that the level of physician-certified absence is higher in IWA-firms than in non-IWA-firms.

the first 16 days of each absence spell is paid for directly by the employer, whereas the social security administration pays for the remaining days and also for subsequent rehabilitation or disability benefits.[4] If a new absence spell starts within 16 days after a previous spell was completed, it is counted as a continuation of the previous spell. This implies that a new pay liability period for the firm is not triggered until the worker has been present for at least 16 days. The social security costs are covered through uniform payroll taxes; hence, there is no experience rating. On average, the pay liability system implies that firms' cover around 34% of the overall sick pay costs for Norwegian workers (see Bjerkedal and Thune 2003). In addition, they bear the administrative costs associated with finding replacements and/or reorganizing the work. Although Norwegian firms are not allowed to cut wages in direct response to an employee's absence behavior, existing empirical evidence indicates that some of the costs are passed on to the employees through an impact on subsequent wage growth (see Markussen 2012).

In April 2002, a reform was implemented implying that firms were entitled to exemption from the 16-days pay liability for pregnancy-related absences. Common symptoms associated with pregnancy-related diseases are nausea, bleeding, anxiety, infections, reduced functional level (e.g., due to back pain), and various psychological reactions. The term "pregnancy related" obviously entails an element of subjective judgment, since pregnant workers also may develop illnesses that have little or nothing to do with their pregnancies. Based on the observed differences in absenteeism between pregnant workers and a comparison group of female nonpregnant colleagues that we present in the next section of this article, we estimate that approximately 73% of the absence among pregnant workers is pregnancy related. We do not observe whether a particular pregnant worker's absence spell is really pregnancy related or not.

During periods of sickness absence, Norwegian workers enjoy a special protection against dismissals, implying that they cannot be dismissed on grounds that are related to their sickness.[5] After the 1-year absence period, however, the firm is allowed to lay off the absent worker with direct reference to the sickness. Hence, if an employer for some reason wishes to lay off a worker—but is prevented from doing so due to the general employment protection regulations—the incentives for facilitating that worker's return to work from a long-term absence spell are particularly weak.

---

[4] Workers who have exhausted their sick pay entitlements but who are still not able to take up work due to sickness are eligible for rehabilitation or disability benefits, depending on the prospects for future recovery/rehabilitation. The replacement ratio for these benefits is around 66%.
[5] The burden of proof lies with the firm. In practice, this implies that absent workers can only be laid off as part of a mass displacement.

Although absence decisions are formally taken by workers and/or physicians, the employer can affect absenteeism in several ways, for example, by (i) monitoring employees (to prevent shirking), (ii) investing in healthy work environments and equipment that can prevent afflictions caused by strain, (iii) offering sick workers modified tasks or changes in the requirements of their jobs, and (iv) allowing the workers more flexible hours and less shift work. These activities obviously involve costs; hence, we may expect employer efforts to depend on the extent to which they bear the direct costs of absenteeism in the first place.

Norway has a generous public program for parental leave benefits. In the period covered by our analysis, there was a paid leave period for 42 weeks with 100% replacement ratio (or 52 weeks with 80% replacement), which has later been raised to 47 weeks (57 weeks with 80% replacement).[6] The leave starts no later than 3 weeks before expected delivery, removing the pregnant workers from the risk of sick leave at this point. It is possible to start the leave period as early as 12 weeks before expected delivery, but this is rarely done in practice.

IV. Data and Empirical Analysis

The data we use in the present article comprise complete longitudinal administrative records on employment and absence for the period 2001–6, merged with information on firms and workers on the basis of encrypted identification numbers.[7] All absence spells are recorded insofar as they are certified by a physician (regardless of recorded symptoms/diagnosis), typically when they exceed 3 days. This implies that the occurrence of very short absence spells is underreported in our data. There is little that we can do about that. To the extent that the removal of pay liability for short-term absenteeism increased the frequency of short absence spells among pregnant workers, we therefore run the risk of underestimating this effect. We return to this issue below. As explained in the previous section, there is also a possibility of self-certifying absence spells as long as 8 days in firms participating in the IWA. Although this option is rarely used in practice (Bergsvik et al. 2010), we deal with the potential difference in certification patterns by always comparing workers who are subject to the same self-certification regulations.

Our empirical analysis consists of two parts. We first examine the extent to which the removal of firms' pay liability for pregnant workers' sick leaves affected these workers absence behavior. We then investigate whether the reform affected young women's employment opportunities. Given that some of the reform effects we seek to evaluate presuppose employers with

------

[6] These replacement ratios apply for annual earnings up to a ceiling of around 500,000 NOK (in 2012 value, approximately US$85,000).

[7] Individual-level absence data before 2001 are not available.

a forward-looking behavior, an appropriate structural model would contain elements of dynamic programming. Our empirical approach is confined to a more reduced-form setting, however, reflecting the more modest aim of evaluating the empirical relevance of direct (myopic) and indirect (forward-looking) responses to changes in firm incentives.

A. Absence Behavior during Pregnancies

To examine the impact of employer incentives on absenteeism we construct the following data set. We start out with all employees who became pregnant between May 2001 and May 2005 and did not make a transition to unemployment, rehabilitation benefits, or disability during the pregnancy. We then follow each pregnant employee for 37 weeks through work presence and sickness absence, that is, until she takes maternity leave and thus is no longer at risk of being absent due to sickness.[8] These records constitute our potential treatment group. Given the rather subtle and potentially conflicting ways in which employer incentives were affected by the removal of pay liability, with likely effects on the incidence as well as the duration of absence spells, we set up a bivariate hazard rate model for transitions between the states of presence and absence to identify the reform effects. But before we turn to that model, we take a closer look at what happened at the exact time of the reform implementation by means of a RD analysis. In both these analyses, we also incorporate a control group of nonpregnant workers; that is, for each pregnancy spell, we pick a female nonpregnant worker from exactly the same workplace, at exactly the same point in time, of approximately the same age, and with a similar earnings level (based on a one-to-one nearest neighbor matching procedure).[9] In the bivariate hazard rate model, the control group plays the essential role of representing the counterfactual trend in absence behavior within a DiD modeling framework. In the RD analysis, we primarily use the control group to perform a placebo analysis.

1. Descriptive Statistics

Table 1 and figures 1 and 2 first offer some descriptive statistics. There are 90,898 pregnancy spells included in our data set and an equally large

---

[8] We assume that the pregnancy started 37 weeks before an observed transition to maternity leave. Transition to maternity leave almost always occurs 3 weeks prior to expected delivery. Only in 1.6% of the cases does the actual delivery take place more than 6 weeks after transition to maternity leave, and these few cases have been deleted from our sample.

[9] We select the coworker with the closest income level, provided that the age difference is less than 3 years. If we cannot find a female coworker with fewer than 3 years of age difference, the pair is not included in our analysis.

Table 1
Descriptive Statistics of Data Used to Analyze Absence Behavior

	Treatment Group (Pregnant)		Control Group (Nonpregnant)	
	Before Reform	After Reform	Before Reform	After Reform
Number of spells	20,845	70,053	20,845	70,053
Characteristics (means):				
Age	30.4	30.7	30.5	30.9
Education:				
Compulsory or lower secondary	11.5	13.1	12.9	15.4
Upper secondary	34.0	34.8	37.7	38.4
College/university	54.6	52.1	49.4	46.2
Earnings (deflated US\$, 2012)	69,130	69,691	66,727	67,467
Non-European background (%)	4.8	4.6	3.8	4.4

number of controls.[10] While our matching procedure is aimed at making the control group as similar as possible to the treatment group, it is evident that some small compositional differences remain. In particular, the pregnant women tend to earn slightly more than their nonpregnant colleagues. This difference can either be attributed to positive selection into pregnancies or to the strong earnings incentives that pregnant workers face, given that their subsequent maternity leave benefits are calculated on the basis of their earnings in the 10-month period just prior to expected delivery. The "pregnancy wage premium" is almost the same before and after the reform, however, approximately 3.5%.

Figure 1 illustrates that the rate of sickness absence has been fairly stable throughout our observation window for both pregnant and nonpregnant female workers, except for seasonal fluctuations. The timing of the reform is marked in the figure as a vertical line. Based on a visual inspection of the figure, it is not easy to spot any reform effect. A closer comparison of absence rates by pregnancy duration before and after the reform is provided in figure 2. Panel *a* clearly illustrates the sharp rise in absence rates that typically occur as the pregnancy progresses, with absence rates as high as 40%–70% the last few months before transition to maternity leave. Panel *b* reports the changes in absence rates from before to after the reform at different stages of the pregnancy for the members of the treatment group minus the corresponding changes for the members of the control group. According to these descriptive DiD estimators, the reform apparently raised absenteeism somewhat, with a possible exception for the later stages of the pregnancy. The overall absence rate for pregnant women rose

---

[10] In total, there were 120,089 pregnancies among employed workers in our data period. We lose 29,191 spells (24.3%) due to lack of appropriate controls.
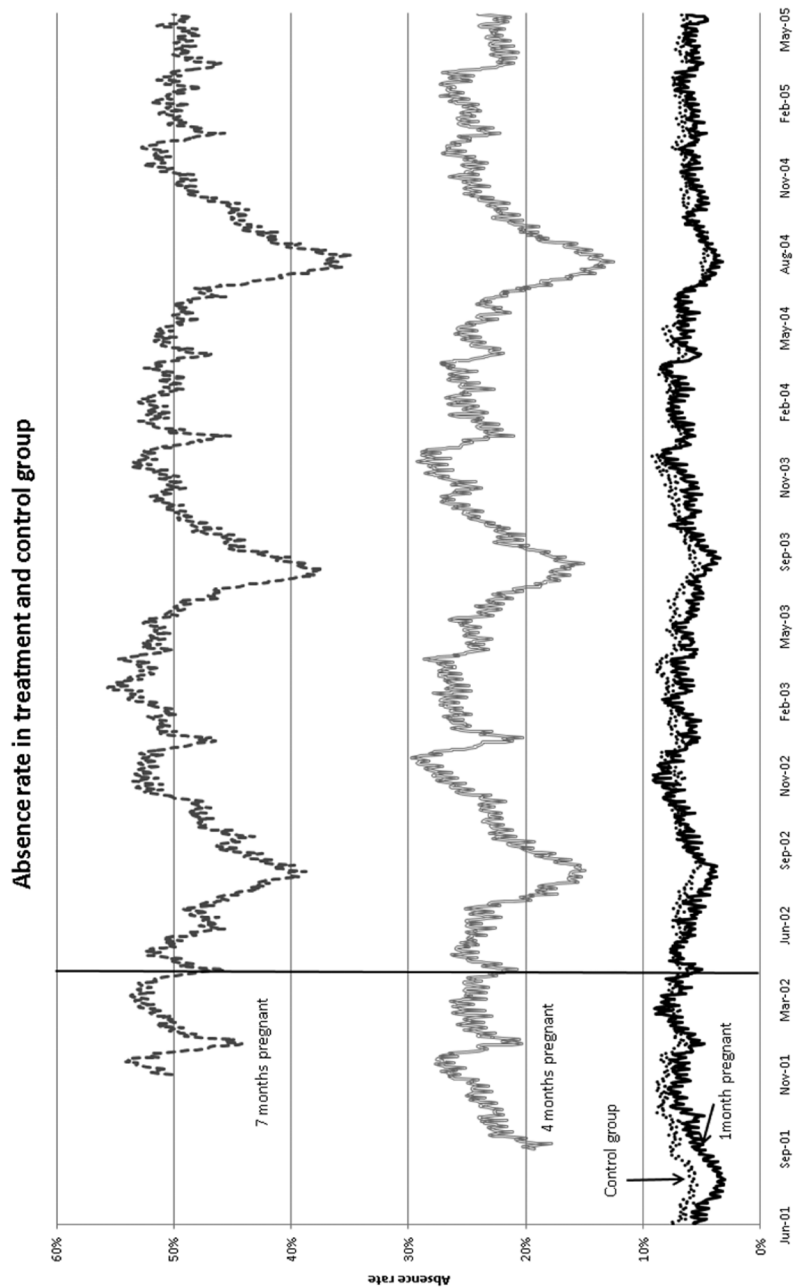
Fig. 1.—Daily absence rates 2001–5 in treatment group (at different stages of the pregnancy) and in control group. Vertical line indicates timing of reform. Daily absence rates are defined as the number of absentees divided by the number of employees at each stage of pregnancy in each of the months displayed.

314
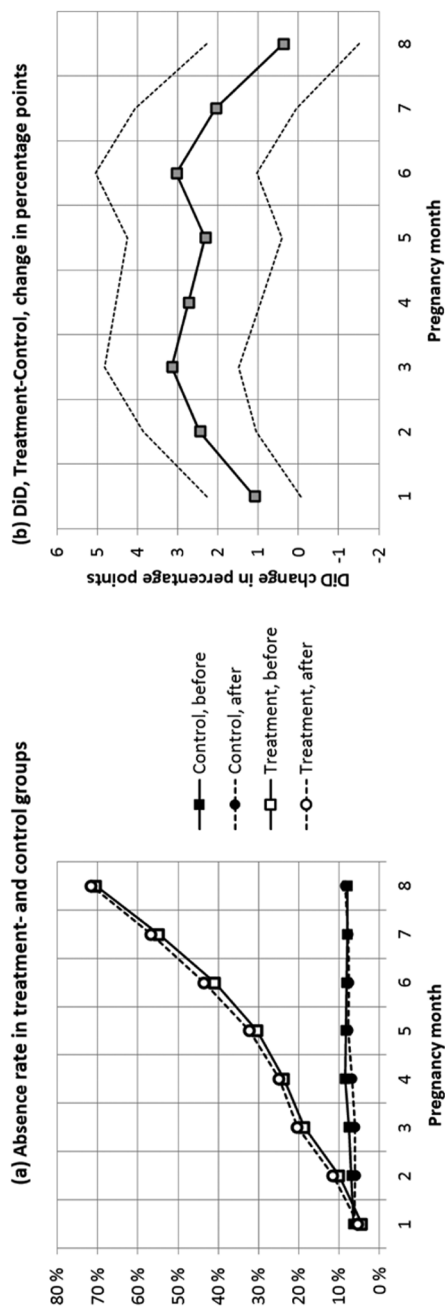
FIG. 2.—Average daily absence rates by number of months into the pregnancy. Treatment and control group before and after the reform (with 95% confidence intervals). For control group members, pregnancy month refers to the number of months since they were picked as controls to a newly pregnant colleague. To prevent seasonal variation from disturbing the before/after reform comparison, the graphs compare workers who became pregnant in May–July 2001 (and their control group members) with those who became pregnant in May–July 2002.

315

from 32.9% before the reform to 33.8% after the reform, while it dropped from 7.5% to 7.3% for the members of the control group. A simple descriptive DiD estimate for the overall reform effect measured by this outcome is thus a modest increase of 1.1 percentage points, or 3.4% (100 × 1.1/32.9).

2. Regression Discontinuity Analysis

Figure 3 illustrates the basic idea of a RD analysis. Panel *a* shows how the daily sick leave entry rates among pregnant workers evolved from 90 days before to 90 days after the reform.[11] Due to the seasonal pattern in absence behavior (recall that the reform was implemented on April 1), there is a declining time trend in this period. Imposing a linear calendar time function to control for this development yields a reform effect corresponding to the small upward shift illustrated in the graph. Panel *c* shows how the size of this estimated reform effect varies as we change the bandwidth (the number of days before/after the reform included in the analysis). The shorter the time window, the larger the estimated effects, suggesting (perhaps) that there was a scope for firms/employees to "postpone" the starting date a few days just around the implementation of the reform. The effect is statistically significant, however, regardless of bandwidth. If we rely on the more conservative estimates, that is, those based on the longest time windows, we conclude (from the size of the jump in panel *a*) that the reform raised the daily entry rate to sick leave among pregnant workers by 0.21 percentage points. Since the expected daily entry rate was estimated to be 1.12% just prior to the jump, this corresponds to an increase of 18.8%. As a sort of "placebo test," we examine in panel *e* how the effect estimate changes if we assume alternative false implementation dates. The graph shows that the effect estimates become smaller the further we move away from the true implementation date. Additional placebo tests can be obtained by repeating the whole estimation procedure on the control group. The results are illustrated in panels *b*, *d*, and *f*. The message coming out of these graphs is clear: there is no reform effect whatsoever in the control group.

A natural objection to the regression discontinuity results presented so far is that they may depend heavily on the linearity assumption for calendar time effects. Table 2 shows, however, that the finding of a positive reform effect is highly robust with respect to the specification of the underlying time function. The table presents results for models with from one to four degree polynomials in a time function assumed to be the same before and after the reform and also with from one to four degree polynomials in time functions estimated separately before and after the reform

[11] Daily entry rates are adjusted for weekdays, public holidays, and the first day after holidays.
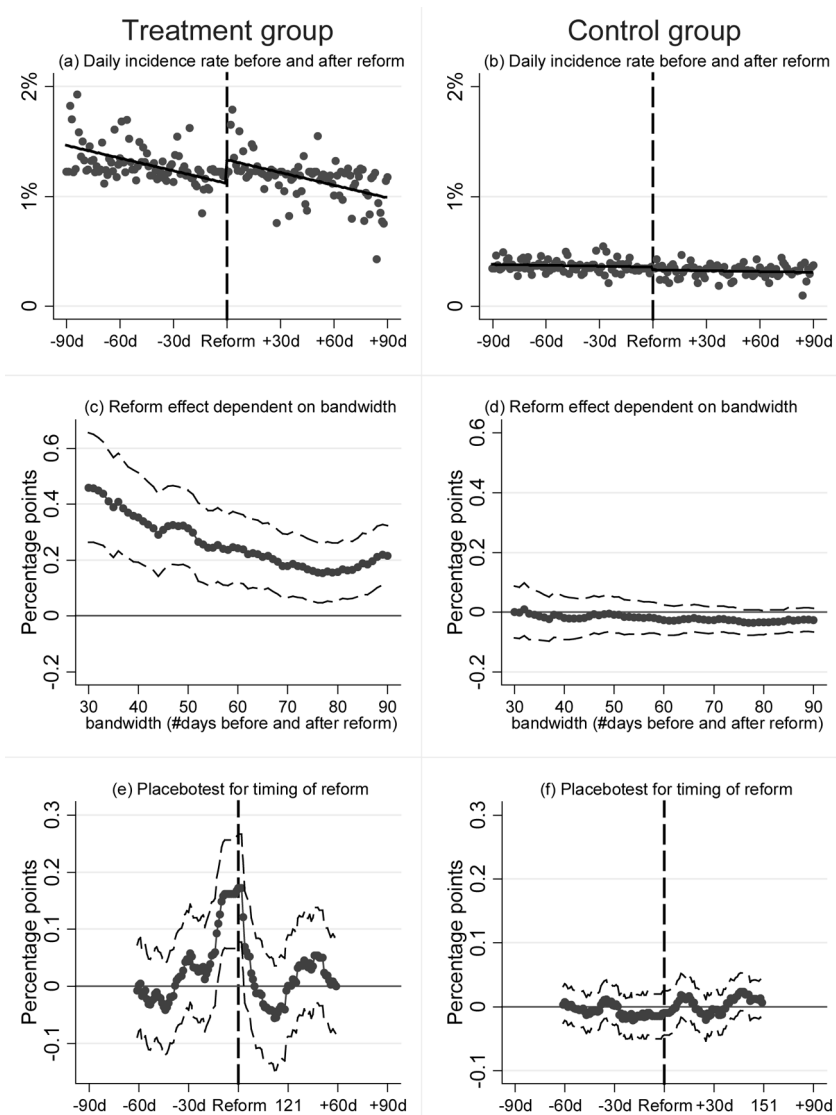
FIG. 3.—Regression discontinuity analysis. Estimated impact on absence incidence of removing firms' pay liability (with 95% confidence intervals). Bandwidth in panels *e* and *f* is 180 days around the true timing of the reform. Reform effects are estimated using a simple linear trend and a reform dummy.

Table 2
Regression Discontinuity Analysis: Estimated Impact on Absence Incidence of Removing Firms' Pay Liability by Bandwidth and Time Function Specification (**Percentage Points Change**)

Polynomial Degree	Common Time Function				Separate Time Functions Before/After			
	1st	2nd	3rd	4th	1st	2nd	3rd	4th
Treatment group:								
Bandwidth (days):								
[−30, 30]	.46	.45	.38	.37	.44	.36	.57	.89
	(4.59)	(4.48)	(2.59)	(2.48)	(4.43)	(2.16)	(2.52)	(2.70)
[−60, 60]	.24	.25	.36	.35	.24	.38	.54	.37
	(3.85)	(3.87)	(4.01)	(3.95)	(3.81)	(3.81)	(3.78)	(2.02)
[−90, 90]	.21	.21	.20	.20	.21	.23	.56	.46
	(3.79)	(3.75)	(2.55)	(2.55)	(3.72)	(2.50)	(4.49)	(2.87)
Control group:								
Bandwidth (days):								
[−30, 30]	.00	−.00	−.02	−.01	.00	−.01	.11	.19
	(.01)	(.00)	(−.32)	(−.18)	(.01)	(−.15)	(1.08)	(1.20)
[60, 60]	−.03	−.03	.00	.00	−.03	.10	−.00	.05
	(−1.10)	(-1.08)	(.08)	(.04)	(−1.07)	(.22)	(−.01)	(.62)
[−90, 90]	−.03	−.03	−.02	−.02	−.03	−.02	.01	.01
	(−1.39)	(−1.33)	(−.80)	(−.83)	(−1.31)	(−.56)	(.30)	(.16)

NOTE.—Values in parentheses are *t*-values.

(in the latter case the reported reform effect is the shift that occurred at the time of reform implementation). All the results are presented for three alternative bandwidths and for both treatments and controls. We interpret the results as convincing evidence that the reform indeed raised the frequency of absence spells among pregnant women. All the 24 alternative specifications produce significant positive reform effects for the treatment group. None of them produce significant results for the control group of nonpregnant colleagues.

Taken at face value, the RD analysis indicates that pregnant workers' entry rate to sick leave was raised 18.8% as a direct result of the reform. If this was the only effect of the reform, the overall level of absenteeism would also increase by 18.8%. This is a huge effect compared to the 3.4% increase suggested by the descriptive DiD calculations referred to above. One possible explanation for this apparent discrepancy is that a higher entry rate was not the only effect but that the reform also triggered significant changes in the duration distribution and recurrence pattern of absence spells and consequently also in the composition of present and absent workers. This is what we examine in the next subsections.

3. A Bivariate Hazard Rate Model

We now turn to a more detailed analysis of the reform effects based on a bivariate hazard rate model where we simultaneously model the transi-

tions to and from sick leave. This simultaneity may be empirically important, since we suspect that the reform affected both transitions, potentially changing the composition of the two risk groups of present and absent workers. Based on economic theory, we hypothesize that four different reform effects may have been at work, all operating through firm incentives:

1. A positive direct effect on absence incidence for workers who had not completed a long-term absence recently (fewer than 16 days ago). This follows directly from the fact that the removal of pay liability made new absence spells less costly for the firm.
2. A negative indirect effect on absence incidence for workers who had completed a long-term absence recently. The reason for this is that the prereform pay liability system entailed a firm incentive to advance any expected (or potential) relapses such that they occurred before a new pay liability period was triggered. The removal of pay liability thus eliminated the incentive to push workers at high risk of relapse into sick leave within 16 days.
3. A negative direct effect on work resumption (positive effect on absenteeism) for absent workers who had been sick-listed fewer than 16 days. This follows directly from the fact that the removal pay liability made absence continuation less costly for the firm within this period.
4. A positive indirect effect on work resumption (negative effect on absenteeism) for absent workers who had been sick-listed more than 16 days. The reason for this is that since the reform removed the pay liability associated with subsequent relapses, it also removed an essential part of the firm's economic risk associated with early work resumption.

An important point to note here is that the indirect effects are not necessarily smaller or less important than the direct effects. We use the term "indirect" to emphasize that the effects in question require an element of forward-looking behavior, not that they are of second order. To the contrary, we expect the indirect work resumption effect (item 4 above) to be particularly significant, as the costs associated with frequent relapses could be very large prior to the reform. Whether the direct or the indirect effects dominate the actual reform responses is an empirical question to which our hazard rate model is designed to provide an answer. Note that the second and third effects listed above are likely to be of minor quantitative importance for overall absenteeism, since they are at work only the first 16 days after or during absence spells, respectively. They are nevertheless of scientific interest, since they convey information about employers' ability to affect the pattern of the employees' absenteeism.

In this exercise, we exploit a much larger time window (2001.5–2005.5) and base our inference on differences in the differences between the treatment and the control groups before and after the reform. Hence, it is the interaction of being pregnant and being in the postreform period that identifies the causal effects of the reform. However, since we are interested in both the direct and indirect effects, the pregnancy-reform interaction needs to be further interacted with variables indicating whether pay liability applied or not in the prereform period, which again depended on the duration of ongoing and past (recent) spells. Since the duration of ongoing and past spells obviously may have autonomous effects on absence behavior in ways that differ between pregnant and nonpregnant workers, we need a full set of controls for these factors to identify the reform effects. Let P be an indicator variable denoting potential treatment (pregnancy), let R be an indicator denoting that the reform has been implemented (after April 1, 2002), let D be an indicator denoting that the current state (presence or absence) has lasted fewer than 16 days, and let I be an indicator denoting that the last absence spell (if any) also lasted fewer than 16 days.[12] Furthermore, let $P_j, j = 9, 8, \ldots, 1$ be indicator variables denoting the number of months until expected delivery for those who are pregnant. We write the hazard rate of moving from presence to absence as

$$
\begin{aligned}
\theta_{1it} = \exp[& x_{it}\beta_{1k} + v_{1i} + \sigma_{1t} + \delta_{1j}P_j \\
& + \phi_{11}D + \phi_{12}D \times R + \phi_{13}D \times P \\
& + \phi_{14}D \times I + \phi_{15}D \times R \times I + \phi_{16}D \times P \times I \\
& + \alpha_{11}P \times R \times \underline{(1 - D + D \times I)} \\
& + \alpha_{12}P \times R \times \underline{(D - D \times I)}],
\end{aligned}
\tag{1}
$$

where $\underline{(1 - D + D \times I)}$ is for "pay liability did apply" and $\underline{(D - D \times I)}$ is for "pay liability did not apply." The hazard rate of moving back to presence (work resumption) is

$$
\begin{aligned}
\theta_{2it} = \exp[& x_{it}\beta_{2k} + v_{2i} + \sigma_{2t} + \delta_{2j}P_j \\
& + \phi_{21}D + \phi_{22}D \times R + \phi_{23}D \times P \\
& + \alpha_{21}P \times R \times \underline{D} + \alpha_{22}P \times R \times \underline{(1 - D)}],
\end{aligned}
\tag{2}
$$

where $\underline{D}$ is for "pay liability did apply" and $\underline{(1 - D)}$ is for "pay liability did not apply." In equations (1) and (2), x_{it} is a vector of observed time-

---

[12] The duration of absence spells is measured inclusive of days belonging to previous absence spells that were terminated fewer than 16 days prior to the start of a current spell.

variant covariates and (v_{1i}, v_{2i}) are unobserved person-specific, time-invariant covariates subject to a joint, but unknown, distribution.[13] The coefficients of interest appear in the last rows of these expressions $(\alpha_{11}, \alpha_{12}, \alpha_{22}, \alpha_{21})$. If employer incentives do affect employees' absence behavior, we expect that $\alpha_{11} > 0$ (effect 1 listed above), $\alpha_{12} < 0$ (effect 2), $\alpha_{21} < 0$ (effect 3), and $\alpha_{22} > 0$ (effect 4). The first rows of equations (1) and (2) contain controls for observed and unobserved heterogeneity, for pregnancies (treatment), and for calendar time. Note that calendar time effects $(\sigma_{1t}, \sigma_{2t})$ are estimated separately for each month, implying that any general differences between the prereform and postreform periods are absorbed by these effects. The second and third of rows of equation (1) and the second row of equation (2) contain controls for the factors that determined initial pay liability, and to ensure that it is the interaction of treatment and reform $(P \times R)$ that identifies the causal effects, these determinants are interacted with P and R separately. This implies that we allow the effects of spell duration (fewer than/more than 16 days) and the effects of recent absenteeism (fewer than/more than 16 days) to vary between the treatment and the control groups and between the prereform and postreform periods.

To avoid setting up a model for the initial state, we condition on workers having been present for at least 32 days to start with.[14] To derive the likelihood function for observed data, we split each individual's event history into parts characterized by constant explanatory variables and unchanged state. Let S_{ki}, $k = 1, 2$, be the set of observed spell parts under risk of event k (sickness, work-resumption) for individual i. Let l_{kis} denote the observed length of each of the spell parts, and let the indicator variables y_{kis} denote whether a spell part at risk of transition k actually ended in such a transition or was right-censored. Conditional on unobserved heterogeneity, the likelihood function for individual i can then be written as

$$L(v_{1i}, v_{2i}) = \prod_{s \in S_{1i}} [\theta_{1it}(v_{1i})]^{y_{1is}} \exp\{-l_{1is}[\theta_{1it}(v_{1i})]\}$$
$$\times \prod_{s \in S_{2i}} [\theta_{2it}(v_{2i})]^{y_{2is}} \exp\{-l_{2is}[\theta_{2it}(v_{2i})]\}. \tag{3}$$

[13] Note that if a person is observed with multiple pregnancies during our observation window, she is treated as multiple persons (with a separate unobserved heterogeneity vector for each pregnancy). The same practice applies if a control person is matched to more than one pregnant colleague. The vector of observed covariates includes age (27 dummy variables, corresponding to ages 19–45), calendar month (56 dummy variables), county (19 dummy variables), income (15 dummy variables), education/industry (15 dummy variables), and the local unemployment rate.

[14] This restriction implies that we drop 2,245 treatment-control pairs (2.5%).

Since the likelihood contribution in equation (3) contains unobserved variables, it cannot be used directly for estimation purposes. This problem may be solved by formulating a model for the distribution of unobserved variables and then replacing equation (3) with its expectation. In the present context, it is probable that the unobserved characteristic affecting absence incidence is correlated to that affecting work resumption, implying that manipulation of the inflow is also likely to change the distribution of unobserved work-resumption propensities in the stock of absentees (and vice versa). We thus model unobserved heterogeneity nonparametrically. Under mild technical assumptions, results in Lindsay (1983, theorem 3.1) and Heckman and Singer (1984, theorem 3.5) ensure that for the purpose of maximizing the likelihood, unobserved heterogeneity may be approximated by a discrete distribution with a finite number of support points. Since only a subset of the analysis population actually experiences an absence spell within our observation window, there is potentially a selection problem involved in estimating the work-resumption hazard. This is similar to the problem arising in the timing-of-events evaluation literature when treatment effects are heterogeneous and only a selected subset of potential participants actually receives treatment. Based on a set of regularity conditions, Abbring and Van den Berg (2003, proposition 4) prove nonparametric identification of the multivariate mixed proportional hazard rate model in this context. To ensure identification in our case, we have performed a Monte Carlo analysis based on artificial data of similar sample sizes and with similar selection challenges that we have in the actual application, that is, data where only a selected portion of the sample comes under the risk of one of the hazard rates during the observation window and with a significant correlation between the two unobserved covariates. The exercise is documented in the appendix, available online. It demonstrates that the sequential hazard rate model we use is indeed identified and that it can be accurately estimated with the algorithm used in this article (with correct standard errors). In our actual application, the foundation for identification is further strengthened by the presence of time-varying covariates as reflected in independent calendar time variation (see McCall 1994; Brinch 2007; Gaure, Røed, and Zhang 2007).

The causal effects of the reform are identified by the shift in pregnant workers' absence behavior from before the reform to after the reform, relative to that of the control group members. An important assumption underlying this identification strategy is that the calendar time effects are the same for the treatment and the control groups. If pregnant workers have been subject to different time trends than nonpregnant workers for reasons that are not related to the reform, the estimated reform effects may be biased. Since we have selected control workers from the pregnant workers' own workplaces, the reform effects may also be biased if colleagues affect each other's absenteeism. Existing evidence indicates that peer ef-

fects in absenteeism are empirically relevant (see, e.g., Ichino and Maggi 2000; Bradley, Green, and Leeves 2007; Hesselius, Nilsson, and Johansson 2009).[15] We return to a number of robustness exercises below, with respect to the composition of the control group, with respect to the assumption of a common trend, and with respect to the modeling of unobserved heterogeneity.

Let Q be the (a priori unknown) number of support points in the unobserved heterogeneity distribution and let $\{(v_{1l}, v_{2l}), p_1\}$, $l = 1, 2, \ldots Q$, be the associated location vectors and probabilities. In terms of observed variables, we write the likelihood function as

$$L = \prod_{i=1}^{N} E_{v_i}[L_i(v_{1i}, v_{2i})] = \prod_{i=1}^{N} \sum_{l=1}^{Q} p_l L_i(v_{1l}, v_{2l}), \quad \sum_{l=1}^{Q} p_l = 1, \quad (4)$$

where $L_i(v_{1i}, v_{2i})$ is given in equation (3). Our estimation procedure is to maximize the likelihood function (4) with respect to all model and heterogeneity parameters repeatedly for alternative values of Q. The nonparametric maximum likelihood estimators (NPMLE) are obtained by starting out with $Q = 1$ and then expanding the model with new support points until the model is "saturated" in the sense that we are no longer able to increase the likelihood function by adding more points. The preferred model is then selected on the basis of the Akaike Information Criterion (AIC). Monte Carlo evidence presented in Gaure et al. (2007) indicates that parameter estimates obtained this way are consistent and approximately normally distributed. They also indicate that the standard errors conditional on the optimal number of support points are valid for the unconditional model as well and hence can be used for standard inference purposes.[16]

Table 3 presents the estimated reform effects and the effects of the key control variables.[17] The two direct effects of the reform contributed to higher absenteeism: the rate of entry into sick leave spells that used to be subject to pay liability rose by around 10% (α_{11}), while the work-resumption rate during the first 16 days of the spells dropped by around 6% (α_{21}). The two indirect effects pulled in the other direction, however: the rate of entry into sick leave spells that were exempted from pay liability already before the reform dropped by as much as 19% (α_{12}), whereas

---

[15] It is also possible that there are peer effects in fertility decisions (see Hensvik and Nilsson 2010). Note, however, that we have removed from the control group all women who themselves become pregnant during the treatment group's pregnancies.

[16] Note that the standard errors do not take any within-group correlation in absenteeism into account, which may be present due to our sampling of control persons within firms. We perform a robustness exercise below, where we have sampled control persons from other firms instead.

[17] A complete listing of estimated coefficients is available from the authors.

Table 3
Selected Estimation Results from Hazard Rate Model

Reform Effect	Coefficient
Reform effect on incidence with initial pay liability (α_{11})	.100***
	(.016)
Reform effect on incidence without initial pay liability (α_{12})	−.215**
	(.119)
Reform effect on work resumption with initial pay liability (α_{21})	−.057**
	(.024)
Reform effect on work-resumption without initial pay liability (α_{22})	.123***
	(.034)
Effects of duration and past absence:	
Effect on incidence of current presence spell lasted fewer than 16 days:	
For all(ϕ_{11})	2.096***
	(.035)
Interacted with reform (ϕ_{12})	.017
	(.030)
Interacted with pregnancy (ϕ_{13})	−1.259***
	(.026)
Additional effect when last absence spell lasted fewer than 16 days:	
For all (ϕ_{14})	.420***
	(.104)
Interacted with reform (ϕ_{15})	−.147
	(.111)
Interacted with pregnancy (ϕ_{16})	.470***
	(.113)
Effect on work resumption of current absence spell lasted fewer than 16 days:	
For all (ϕ_{21})	.978***
	(.031)
Interacted with reform (ϕ_{22})	.108***
	(.032)
Interacted with pregnancy (ϕ_{23})	.253***
	(.037)

NOTE.—Number of observations = 177,306. Standard errors are in parentheses. Additional controls include age (34 dummy variables), calendar time (64 dummy variables), county (19 dummy variables), income (15 dummy variables), education/industry (15 dummy variables), and local unemployment (in the municipality). The model also includes six support points for the two-dimensional unobserved heterogeneity distribution and the estimated correlation coefficient between $\exp(v_{1i})$ and $\exp(v_{2i})$ is .30. In total, the model contains 348 estimated parameters.
** Significant at the 5% level.
*** Significant at the 1% level.

the work-resumption hazard from long-term absences rose by 13% (α_{22}).[18] All these results are exactly as one would expect on the basis of economic theory. Firms responded to the new incentives by reducing efforts to prevent short-term absence, for which they no longer faced any direct costs. At the same time, they apparently became less skeptical toward allowing

[18] The estimated percentage shift in the hazard rates caused by a given variable is given as 100 × (exp(estimate) −1), which for small estimates is approximately equal to the estimate multiplied by 100.

long-term absentees back into work and became more willing to let "risky" presence spells exceed the limit of 16 days, knowing that they no longer were responsible for the sick pay costs associated with relapses.

In order to examine the quantitative importance of short-term pay liability, we have performed a simulation exercise based on the estimated model; that is, we have simulated a large number of spells for typical pregnant workers with and without the direct and indirect effects reported in table 3 "turned on."[19] When we only allow the two direct effects $(\alpha_{11}, \alpha_{21})$ to enter the model (i.e., disregard the indirect effects), it predicts that pay liability reduces the overall expected number of absence days during pregnancies from 85.5 to 79.3, or by 7.3%. However, when we also allow the two indirect effects $(\alpha_{12}, \alpha_{22})$ to enter the model, the predicted number of absence days again rise to 82.8, reducing the overall effect of pay liability to a 3.2% reduction. It is instructive to decompose the predicted effects of short-term pay liability into absence days subject to and absence days exempted from pay liability. Based on our simulations, we find that while pay liability for short-term absences reduces the number of absence days directly covered by pay liability by around 6.8%, it reduces the number of noncovered days by 3.5% only. If we extrapolate our findings to nonpregnant workers, we find that pay liability on short-term absenteeism reduces short-term absenteeism by as much as 11.1%, while it actually increases long-term absenteeism by around 1.6%. The reason why we predict that the effects of pay liability are relatively more important for nonpregnant than for pregnant workers is that a disproportionally large fraction of pregnant workers are "stuck" in long-term absenteeism anyway, with return-to-work rates return so low that the negative proportional shift caused by pay liability (α_{22}) is of moderate quantitative importance. Consequently, they also become removed from the risk of incidence.

It is important to bear in mind that all the effect estimates reported here are likely to underestimate the true causal effect of pay liability for the reason that the reform did not really affect all absences among pregnant workers—only those that were "pregnancy related." In Section IV.A.3 above, we estimated that only 73% of pregnant workers' absence spells are pregnancy related, implying that our coefficients are estimated with a significant attenuation bias.

4. Robustness

As discussed above, the results in table 3 are based on the assumption that young, nonpregnant, female colleagues constitute an appropriate con-

---

[19] Simulations were made for an agent starting a pregnancy with "representative" entry and exit rates, based on the point estimates reported in table 3. We made 177,176 simulations for each set of assumptions regarding the reform effects (corresponding to the actual number of observations).

trol group for pregnant workers, in the sense that they were subject to the same time variation in absence behavior apart from the reform and also were not affected by the reform through peer effects. Yet, while it is conceivable that violations of these assumptions could disturb attempts to identify overall reform effects, it is hard to see why they should yield the particular twist in the work-resumption profiles captured by our estimates. We interpret the finding that the work-resumption rate declined during the first 16 days of absence (for which there used to be a pay liability period) while at the same time it rose at longer durations as convincing evidence of a reform effect.

Table 4 reports the results from a number of robustness exercises. We first examine sensitivity with respect to the common trend assumption. Column 2 reports the estimated reform effects when the time window is reduced to include 22 months only (dropping the last 42 months), while columns 3 and 4 report the results when instead a separate linear or quadratic time trend for pregnant workers is included in the model, respectively. While these modifications do change the estimated effects somewhat—in the direction of weaker direct effects and (in some cases) stronger indirect effects—they do not alter any of the main conclusions. The sensitivity of the point estimates with respect to the common trend assumption may suggest, however, that there was an increasing trend in pregnancy-related absences for reasons not related to the reform. Alternatively, it may have been the case that the full reform effects materialized gradually, as it took some time before employers learned about the new rules. Second, we assess the potential problem that the absence behavior in the control group is causally affected by the absence behavior in the treatment group. Since our control persons are chosen from the same workplaces as the treatment population (to ensure maximum similarity in terms of shocks to the work environment), such a causal relationship could arise due to peer effects or changes in work pressure resulting from colleagues' absence. Column 5 reports the results from a model where we have drawn the control group members from other workplaces instead; that is, for each treatment, we have selected a control person from a different firm within the same industry and with the same IWA status but have otherwise followed the same one-to-one nearest neighbor matching strategy as described above.[20] It is clear that the estimated direct incidence effect as well as the indirect effect on work resumption then become larger. This pattern is consistent with the existence of a significant peer effect, in line with existing empirical evidence. Again, the main conclusions go through, however.

In our baseline model, we have imposed rather restrictive assumptions regarding duration dependence in the return-to-work hazard by allowing

---

[20] Industries are categorized on the basis of the International Standard Industrial Classification (ISIC) at a three-digit level.

Table 4
Robustness: Estimated Reform Effects under Alternative Modeling Assumptions

	Baseline Model (Repeated from Table 3) (1)	Reduced Time Window (2001.5–2003.3) (2)	Separate Linear Time Trend for Pregnant (3)	Separate Quadratic Time Trend for Pregnant (4)	Control Group Selected from Other Workplaces (5)	With Flexible Duration Baseline (Nine Steps) (6)	Without Unobserved Heterogeneity (7)	Absence-Specific Unobserved Heterogeneity (8)
Reform effect on incidence with initial pay liability (α_{11})	.100*** (.016)	.082*** (.019)	.042** (.020)	.066** (.029)	.141*** (.015)	.100*** (.015)	.096*** (.014)	.105*** (.018)
Reform effect on incidence without initial pay liability (α_{12})	−.215** (.119)	−.226 (.143)	−.276*** (.117)	−.256** (.121)	−.094 (.112)	−.224** (.117)	−.221** (.102)	−.268** (.103)
Reform effect on work-resumption with initial pay liability (α_{21})	−.057** (.024)	−.031 (.028)	−.049* (.029)	.034 (.039)	−.020 (.022)	−.076*** (.024)	−.058** (.023)	−.048** (.023)
Reform effect on work-resumption without initial pay liability (α_{22})	.123*** (.034)	.107** (.041)	.133*** (.038)	.208*** (.045)	.174*** (.031)	.109*** (.034)	.124*** (.033)	.106*** (.033)
Number of support points in unobserved heterogeneity distribution	6	4	5	3	2	5	1	10

NOTE.—Standard errors are in parentheses.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

only a two-step function in the form of a shift after 16 days of absence. To assess whether this may have influenced our results, we report in column 6 estimates based on a much more flexible nine-step function (allowing a shift every thirtieth day). Again, we find that the results are robust, although we estimate a somewhat stronger reform response during the first 16 days and correspondingly weaker response afterward. Finally, we examine robustness with respect to the modeling of unobserved heterogeneity. Column 7 reports the results from a model without unobserved heterogeneity, while column 8 reports the results from a model where the unobserved heterogeneity vector is allowed to change each time a person has returned to work from a completed absence spell (and been at work for at least 16 days). The latter specification is included to allow distinct work-absence cycles to be associated with different unobserved transition propensities, generated, for example, by different diseases. The results change only marginally, indicating that the exact way in which unobserved heterogeneity is modeled is empirically unimportant in this case.

5. Comparison with RD Findings and Descriptive Statistics

While the purely descriptive DiD estimate for the overall reform effect was a 3.4% rise in absenteeism (see Sec. IV.A.1), the simulation exercise based on the hazard rate model indicates a rise around 3.2%. These two very different empirical approaches thus yield more or less the same conclusion regarding the overall effects of the reform. The RD analysis, on the other hand, indicated a shift of the incidence rate alone of 18.8%. This is not necessarily in conflict with the results from the DiD and the hazard rate approaches, since the latter also incorporate effects on spell duration. But even for the incidence effect, the hazard rate model and the RD analysis seem to come up with different results. While the RD analysis indicated a shift in the incidence rate 18.8%, the hazard rate model indicates a shift around 10%. This difference can be explained, however, by endogenous sorting into the risk set of present workers, which is appropriately accounted for in the bivariate hazard rate model but not in the RD analysis. To show this, we reestimated the hazard rate model without including controls for pregnancy-month (and other individual characteristics) in the model, making it more directly comparable to the RD analysis (not shown in tables). We then obtained an entry effect (α_{11}) of 0.18 (SE = .01), which is basically in line with the RD estimates with the longest bandwidths. Hence, what apparently happened was that the reform caused entry rates to rise for two very different reasons: (i) because transition to absenteeism for each present worker became less costly for the firm and (ii) because the composition of the population of present workers changed toward individuals with higher absence risks (as long-term absentees to a larger extent returned to

work). While the bivariate hazard rate model is designed to disentangle these two effects, the RD analysis lumps them together. This illustrates a potential limitation associated with using RD in settings were the source of the discontinuity affects behavior such that the composition of the agents under study also changes; the method may be a reliable and robust strategy for ascertaining that behavior really has changed yet be a poor strategy for quantifying these effects at the level of agents.

B. Employment Opportunities for Young Females

The firms' pay liability for pregnant workers' absences was removed for the purpose of making individuals conceived to have a high risk of becoming pregnant more employable. In this subsection, we evaluate whether the reform had this intended effect or not. Employers' scope for discriminating against pregnant or pregnancy-prone workers is limited insofar as workers already are employed. Hence, to the extent that the higher expected absence costs associated with pregnancies affect employment opportunities at all, they are likely to do so primarily through the hiring decision. We therefore start out with the population of labor-market entrants, that is, persons who completed their educational career (upper secondary or higher) in the period from 2000 to 2004, and we investigate transitions to a first job.[21] It is not obvious how a treatment group should be defined in this case, and it is even less obvious how an appropriate control group can be established. Loosely speaking, the treatment group consists of women at risk of becoming pregnant, which from the employers' point of view basically includes all young women. In a more narrow sense, it consists of already-pregnant job seekers. As our primary strategy, we thus define all young, female labor-market entrants (aged 19–34) to be in the treatment group, whereas we use men in the same age group as controls. In addition, we examine in particular the employment prospects of pregnant graduates.

We first show some descriptive statistics illustrating the school-to-work transitions for males and females prior to and after the reform. We then set up a multivariate regression model accounting for the transition to the first job and investigate whether the reform coincided with a relative improvement in employment chances for female labor-market entrants in general and for pregnant entrants in particular.

1. Descriptive Statistics

Descriptive statistics for the treatment and control groups are provided in figure 4 and in table 5. We focus on groups who either completed their

---

[21] A person is considered to have completed education in a particular quarter if he or she studied in that quarter but was not registered in any formal education the subsequent two quarters.
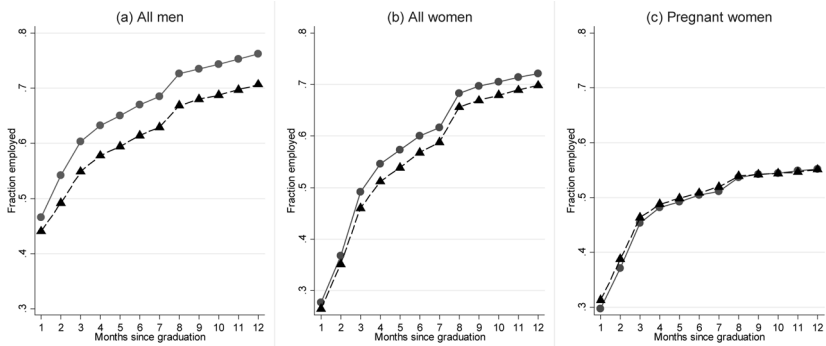
F$_{IG}$. 4.—Cumulative employment fractions by months since graduation. Before reform, solid gray line with circles; after reform, dashed black line with triangles.

education well before (June 2000–March 2001) or well after (June 2002–March 2003) the reform to avoid cohorts that were only partly affected by the reform. Looking first at the cumulative employment propensities by time since graduation in figure 4, we note that both women and men experienced a significant drop in employment from the former to the latter period but that the drop was largest for men, implying a relative improvement for women (see panels *a* and *b*).[22] Moreover, women who were pregnant at the time of graduation did not experience a drop in employment propensity at all, thus experiencing a quite substantial relative improvement. The descriptive DiD estimator for the reform's effect on young women's overall probability of having found some form of employment within 6 or 12 months after graduation is 2.4 and 3.2 percentage points, respectively (see the numbers reported at the bottom of table 5).

Table 5 also reveals that men and women tend to chose very different educations, as reflected in the much higher level of college and university education for women than for men. In addition, it reveals that a much larger fraction of men than of women has obtained employment well before the time of graduation. An important reason for this is that many secondary educations involve the possibility of a final apprentice year, which implies employment.

2. A Probability Model for Subsequent Employment

Given the large differences in men's and women's educational choices, it is also probable that their school-to-work transitions have been subject to correspondingly different cyclical and seasonal fluctuations around the time of the reform. To avoid this from biasing our results, we set up a

---

[22] Employment is defined on the basis of the Norwegian Employer-Employee register. A person is defined to have been employed in a particular month if he or she had a registered job paying a monthly salary of at least NOK 20,500 (2012 value).

Table 5
Descriptive Statistics of Data Used to Analyze Transitions to Employment

	Treatment: All Young Women		Control: All Young Men	
	Before June 2000– March 2001	After June 2002– March 2003	Before June 20000– March 200101	After June 2002– March 2003
Number of graduates	19,589	18,843	19,674	18,462
Age	25.5	25.3	25.1	24.7
Education				
Upper secondary	44.7	46.8	64.0	65.7
College (low level)	43.6	41.0	22.2	20.7
University (high level)	11.7	12.2	13.9	13.6
Non-European background (%)	5.2	6.1	5.4	5.7
Cumulative fraction employed (%):				
6 months before graduation	14.7	13.9	35.9	33.2
At graduation	19.6	19.4	39.9	38.3
6 months after	60.0	56.8	67.0	61.4
12 months after	72.1	69.8	76.2	70.7

multivariate regression model for the school-to-work transition, where we control for all possible combinations of education types and graduation times; that is, we define 49 different education categories and interact them with dummy variables for the before/after reform period.[23]

We focus on the two discrete outcomes of having obtained some form of employment within 6 or 12 months after graduation. For ease of interpretation, we specify the school-to-work transition with a linear probability model, that is,

$$E[y_i] = x_i\beta_3 + \phi_{31}\text{EDUC}_i + \phi_{32}\text{EDUC}_i \times R + \phi_{33}W + \phi_{34}P \\ + \alpha_{31}W \times R + \alpha_{32}P \times R, \tag{5}$$

where y_i is an indicator for employment within 6 or 12 months, W is a dummy for women, and EDUC_i is a vector with 49 education dummy variables. The pregnancy dummy P is equal to one if a woman was pregnant at the time of graduation or within the first 3 months after graduation. The covariate vector x_i includes the month of graduation (17 dummy variables), age (16 dummy variables), and the local (municipality) rate of youth unemployment (aged 19–34) at the time of graduation. The coefficients of main interest are α_{31} and α_{32}.

---

[23] Educations are classified on the basis of the International Standard Classification of Education (ISCED97).

The main regression results are presented in columns 1 and 2 of table 6. For the 12-month outcome, we find a small positive reform effect on young women's employment propensity equal to 1.6 percentage points. The effect is statistically different from zero at the 5% level. For women who were pregnant at the time of graduation, we find an additional effect of 2.9 percentage points, significant at the 10% level. The effects are imprecisely estimated, however. When we focus on a 6-month period rather than a 12-month period, the estimated effect for women as a whole declines (and become statistically insignificant), whereas the additional effect for pregnant women becomes stronger.

As shown in table 5, many students—particularly men—have obtained some employment long before graduation. Since it is difficult to interpret the school-to-work transition for these graduates, we have also estimated the model without including those who had a job as long as 6 months before graduation. The results from this exercise are displayed in columns 3 and 4 of table 6. While the reform effects for women as a whole are virtually unaffected by this sample restriction, the estimated effects for pregnant women become somewhat smaller. To sum up, it seems that the removal of pay liability for pregnancy-related absences had the intended

Table 6
Estimated Reform Effects on Employment Propensity

	Whole Sample		Reduced Sample: No Employment 6 Months before Graduation	
	Employment within 12 Months (1)	Employment within 6 Months (2)	Employment within 12 Months (3)	Employment within 6 Months (4)
Effect for all women (α_{31})	1.59**	1.00	1.62*	1.08
	(.70)	(.76)	(.84)	(.89)
Effect for pregnant (α_{32})	2.84*	4.13**	2.03	3.50*
	(1.51)	(1.65)	(1.80)	(1.91)
Women (ϕ_{33})	−.35	−3.00***	2.29***	.02
	(.49)	(.53)	(.59)	(.63)
Pregnant (ϕ_{34})	−19.87***	−12.09***	−22.48***	−13.98***
	(1.04)	(1.14)	(1.24)	(1.32)
Local youth unemployment rate (ages 19–34)	−6.89***	−7.65***	−6.78***	−7.24***
	(.05)	(.06)	(.06)	(.07)
Observations	76,568	76,568	57,864	57,864

NOTE.—Additional controls include age (16 dummy variables), graduation month/year (17 dummy variables), and education type/level before/after reform (97 dummy variables).
* Significant at the the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

effect of making young women more employable. Our best guess is that it raised the probability of having experienced some form of employment 1 year after graduation by 1–2 percentage points for women in general and by 4–6 percentage points for pregnant graduates. To put these estimates into perspective, it may be of interest to examine the sizes of the expected sick pay costs that were effectively removed by the reform. Based on the simulation exercise described in the previous subsection, we compute that being pregnant implies approximately 12 extra sick pay days directly covered by pay liability. The reform's removal of this cost corresponds to a 5% reduction of the total wage bill during a typical pregnancy. On average, female graduates aged 23–27 experienced 0.4 pregnancies during the first 3-year period after graduation. Hence, the reform did have a small, but noticeable, effect on expected wage costs for new female employees.

V. Conclusion

The findings reported in this article show that employees' sick leave behavior is responsive toward the employer's wage costs during their absence. If a significant part of the costs can be passed on to a public insurer, employees tend to be absent more than if the costs are paid for by their own employer. In most countries, the employer is responsible for the costs associated with short-term absence, while the public insurer covers the direct sick pay costs arising from long-term absence. We have shown that an insurance system with these properties reduces the incidence of absence spells and also raises the probability of quick work resumption. At the same time, however, it also reduces the work-resumption rate at durations exceeding the pay liability period. We conclude that responsibility for short-term sick pay only undermines the firms' incentives to prevent long-term absenteeism, not only because they have too little pecuniary incentive to avoid long-term absence per se but also because a long-term absent worker's return to work entails the risk of costly short-term relapses (for which the employer is again financially responsible). As a result, employers may exert too little effort to prevent long-term absence. Although we must be careful generalizing the quantitative results obtained for pregnant workers to workers at large, we see no reason why the phenomenon discovered in this article should be restricted to pregnant employees. The evidence presented in this article then indicates that the unintended side effect of restricting pay liability to short-term absence spells only—in terms of raising longer-term absenteeism—may be sufficiently large to almost nullify the favorable impacts of a 16-day pay liability period for short-term absenteeism. Extrapolating our findings to the work force at large, we estimate that by making the employers responsible for around one-third of overall sick pay costs in Norway, policy makers achieve a mere 2.7% reduction in overall absence and even a slight increase in long-term absence.

Given that long-term absence is the typical gateway to permanent disability benefits (which are fully covered by taxpayers), insufficient employer incentives to prevent long-term absenteeism may be very costly from a social point of view.

Why, then, have so many countries designed their sickness insurance systems such that employers do not face the direct sick pay costs associated with long-term absenteeism? We see two possible explanations. The first is that absence behavior has typically been considered to be determined by the worker—not the firm. Hence, the focus has been placed on worker incentives rather than firm incentives. The results presented in this article show that this argument is not valid; firms do influence their employees' absence behavior, and they respond to economic incentives. Second, there may be significant administrative costs associated with reimbursing firms for the large number of short-term absences, implying that even a modest pay liability for longer absence spells is difficult to implement without at the same time raising the overall sick leave costs for firms. While this may be a desirable outcome from the perspective of absence prevention, it has been argued that it also undermines employment prospects for job seekers with high expected absence propensity. Our findings indicate that, to some extent, this argument is valid; firms do respond to sick pay liability by being less willing to employ workers expected to be absent a lot. Hence, by raising pay liability for long-term absences without at the same time reducing it for shorter spells, policy makers may undermine incentives to employ marginal workers. A possible solution to this dilemma may be to restructure firms' pay liability so that firms cover a smaller fraction of the short-term absence costs and a larger fraction of the long-term absence costs. This solution requires, however, that the administrative challenges associated with high numbers of reimbursement transactions can be overcome.

We conclude this article with a caveat. Since our data include physician-certified absences only, we obviously cannot evaluate the impacts of pay liability on self-reported absenteeism. While we see no particular reason to expect significant differences in the responses of self-certified and physician-certified sick leave to changes in firm incentives, we cannot rule out that such differences exist.

References

Abbring, Jaap H., and Gerard J. Van den Berg. 2003. The nonparametric identification of treatment effects in duration models. *Econometrica* 71, no. 5:1491–1517.

Bergsvik, Daniel, Simen Markussen, and Oddbjørn Raaum. 2010. Tok en tredagers? Mønstre i egenmeldt sykefravær (Three days off? Patterns of self-reported absenteeism). *Søkelys på Arbeidslivet* 27, no. 4:379–95.

Bjerkedal, Tor, and Ola Thune. 2003. Hva koster sykelønnsordningen? (Total cost of sick leave benefits: The case of Norway). *Tidsskrift for den Norske Lægeforening* 123 (May): 662–63.

Bonato, Lusine, and Leo Lusinyan. 2007. Work absence in Europe. *IMF Staff Papers* 54, no. 3:475–538.

Bradley, Steve, Colin Green, and Gareth Leeves. 2007. Worker absence and shirking: Evidence from matched teacher-school data. *Labour Economics* 14, no. 3:319–34.

Brinch, Christian. 2007. Nonparametric identification of the mixed hazards model with time varying covariates. *Econometric Theory* 23, no. 2: 349–54.

Burkhauser, Richard V., and Mary C. Daly. 2011. *The declining work and welfare of people with disabilities: What went wrong and a strategy for change*. Washington, DC: AEI.

Burkhauser, Richard V., Maximilian D. Schmeiser, and Robert R. Weathers. 2012. The importance of anti-discrimination and workers' compensation laws on the provision of workplace accommodations following the onset of a disability. *Industrial and Labor Relations Review* 65, no. 1: 350–76.

Cahuc, Pierre, and André Zylberberg. 2004. *Labor economics*. Cambridge, MA: MIT Press.

D'Amuri, Francesco. 2011. Monitoring, monetary incentives, and workers' rents in determining absenteeism. Working paper, Research Department, Italian Central Bank. http://sites.google.com/site/fradamuri /home/research.

Edwards, Paul, and Kay Greasley. 2010. *Absence from work*. Report from the European Foundation for the Improvement of Living and Working Conditions. Dublin: European Working Conditions Observatory.

Gaure, Simen, Knut Røed, and Tao Zhang. 2007. Time and causality: A Monte Carlo evaluation of the timing-of-events approach. *Journal of Econometrics* 141, no. 2:1159–95.

Gruber, Jonathan. 1994. The incidence of mandated maternity benefits. *American Economic Review* 84, no. 3:622–41.

Heckman, James, and Burton Singer. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52, no. 2:271–320.

Henrekson, Magnus, and Mats Persson. 2004. The effects on sick leave of changes in the sickness insurance system. *Journal of Labor Economics* 22, no. 1:87–113.

Hensvik, Lena, and Peter Nilsson. 2010. Businesses, buddies and babies: Social ties and fertility at work. IFAU Working Paper 2010:9, Institute for Evaluation of Labour Market and Education Policy (IFAU), Uppsala, Sweden.

Hesselius, Patrik, J. Peter Nilsson, and Per Johansson. 2009. Sick of your colleagues' absence? *Journal of the European Economic Association* 7, nos. 2–3:583–94.

Ichino, Andrea, and Giovanni Maggi. 2000. Work environment and individual background: Explaining regional shirking differentials in a large Italian firm. *Quarterly Journal of Economics* 115, no. 3:1057–90.

Johansson, Per, and Mårten Palme. 2005. Moral hazard and sickness insurance. *Journal of Public Economics* 89, nos. 9–10:1879–90.

Lindsay, Bruce G. 1983. The geometry of mixture likelihoods: A general theory. *Annals of Statistics* 11:86–94.

Markussen, Simen. 2012. The individual cost of sick leave. *Journal of Population Economics* 25, no. 4:1287–1306.

Markussen, Simen, Knut Røed, Ole J. Røgeberg, and Simen Gaure. 2011. The anatomy of absenteeism. *Journal of Health Economics* 30, no. 2: 277–92.

McCall, Brian P. 1994. Identifying state dependence in duration models. In *Proceedings of the Business and Economics Section, American Statistical Association, 1994,* 14–17. Washington, DC: American Statistical Association.

OECD (Organization for Economic Cooperation and Development). 2010. *Sickness, disability and work: Breaking the barriers*. Paris: OECD.

Petterson-Lidbom, Per, and Peter Skogman Thoursie. 2013. Temporary disability insurance and labor supply: Evidence from a natural experiment. *Scandinavian Journal of Economics* 115, no. 2:485–507.

Shavell, Steven, and Laurence Weiss. 1979. The optimal payment of unemployment benefits over time. *Journal of Political Economy* 87, no. 6: 1347–62.

Topel, Robert H. 1983. On layoffs and unemployment insurance. *American Economic Review* 73, no. 4:541–59.

Ziebarth, Nicolas. 2009. Long-term absenteeism and moral hazard: Evidence from a natural experiment. Working Paper 888, DIW Berlin.

Ziebarth, Nicolas, and Martin Karlsson. 2009. The effects of expanding the generosity of the statutory sickness insurance system. Health, Econometrics, and Data Group (HEDG) Working Papers 09/35, HEDG, University of York.