Does Menstruation Explain Gender Gaps in Work Absenteeism?

Jonah E. Rockoff

Mariesa A. Herrmann

April, 2011*

Abstract

Ichino and Moretti (2009) find that menstruation may contribute to gender gaps in absenteeism and earnings, based on evidence that absences of young female Italian bank employees follow a 28-day cycle. We find this evidence is not robust to the correction of coding errors or small changes in specification, and we find no evidence of increased female absenteeism on 28-day cycles in data on school teachers. We show that five day workweeks can cause misleading group differences in absence hazards at multiples of seven, including 28 days, and illustrate this problem by comparing absence patterns of younger males to older males.

---

* Jonah Rockoff is the Sidney Taurel Associate Professor of Business at Columbia Business School and a Research Associate at the National Bureau of Economic Research. Mariesa Herrmann is a doctoral candidate in the Columbia University Department of Economics. Correspondence should be sent to jonah.rockoff@columbia.edu. The authors thank Andrea Ichino and Enrico Moretti for sharing their data and computer code and for helpful discussions. They also thank Doug Almond, Janet Currie, Lena Edlund, Ray Fisman, Wojciech Kopcuk, Cecilia Machado, Bentley MacLeod, Alexei Onatski, Doug Staiger, and the seminar participants at the Applied Micro Colloquium and Econometrics Colloquium at Columbia University for their thoughtful comments and suggestions. Mariesa Herrmann gratefully acknowledges the financial support of a National Science Foundation Graduate Research Fellowship. Any opinions, findings, conclusions or recommendations expressed in this study are those of the authors and do not necessarily reflect the views of the National Science Foundation. Access to the data used in this article is limited to protect confidentiality, but can be obtained beginning six months after publication through three years hence; the Italian Bank data can be obtained from Andrea Ichino, Piazza Scaravilli 2, 40126 Bologna, Italy, andrea.ichino@unibo.it; the New York City teacher data can be obtained from Jonah Rockoff, Uris Hall, 3022 Broadway, New York, NY 10027-6902, jonah.rockoff@columbia.edu.

I. Introduction

A large literature in economics documents differences in earnings between men and women (see Goldin 1990, Blau and Kahn 2000). In a recent paper, Ichino and Moretti (2009), hereafter IM, argue that gender gaps in earnings can be partially explained by increased female absenteeism related to menstruation. While the higher rate of absenteeism among women is a well-known fact (see Paringer 1983), IM present new evidence on the role of menstruation using data from a large Italian bank, where they find that, relative to same-aged men, women under the age of 45 exhibit a high rate of absence spells initiating at 28-day intervals but women over the age of 45 do not.

Standard explanations for the gender earnings gap include gender differences in preferences, gender differences in skills, and discrimination (Altonji and Blank 1999). The notion that biological differences between men and women partially explain gender gaps in labor market outcomes is novel, provocative, and deserving of scrutiny. Indeed, IM are careful to note that their "findings are based on data from only one firm and their external validity is unclear."

This paper reexamines the evidence and finds no support for the notion that menstruation is an important determinant of gender gaps in absences and earnings. First, we highlight an econometric issue related to the five-day workweek that can cause misleading group differences in the hazard rates of absence spells initiating at intervals that are multiples of seven, including 28. Second, we revisit the study of absenteeism among Italian bank employees.[1] We demonstrate that the estimates suggesting a significantly higher hazard rate of absence spells initiating at 28 days for younger women at 28 days are sensitive to the correction of coding errors and allowing

[1] We thank Ichino and Moretti for generously sharing their data and computer code, as well as answering a number of our questions and discussing our findings.

for serial correlation within individuals. We also illustrate how the distortions induced by the five-day workweek cause large, significant differences in the hazard rates at 28 days (and other multiples of seven) between two groups who do not experience menstruation: younger and older men. The difference between the hazard rates of younger and older men at 28 days is at least as large as that between younger men and women, cautioning against attributing differences in absence hazard rates at 28-days to menstruation. Finally, we conclude with a similar analysis using data on New York City public school teachers, which provides an informative comparison with the Italian bank data, given the many differences between the two institutional settings.[2] While female teachers are absent more often than their male colleagues, we find no evidence that the initiation of absence spells for younger female teachers follow 28-day cycles. However, as this analysis also suffers from the distortions induced by the five-day workweek, we do not claim that this is conclusive evidence that menstruation is not a significant determinant of absenteeism for teachers, who comprise a substantial segment of the female labor force in the U.S.[3]

The rest of the paper proceeds as follows. Section II describes the econometric model and the importance of the five-day workweek. Section III presents our re-analysis of the Italian bank data, and Section IV describes the New York City data and our analysis. Section V concludes.

---

[2] In contrast to female Italian bank employees, teachers are highly educated and may face higher financial and psychic costs of absence (for instance because they care about student achievement), and Americans may face different cultural norms about work absence due to menstruation, or work absence more generally.

[3] In 2008, teaching in an elementary or middle school was the third most prevalent occupation for employed women in the United States, behind secretaries/administrative assistants and registered nurses; 3.5 percent of employed women worked as elementary or middle school teachers (http://www.dol.gov/wb/stats/main.htm, accessed 5/16/2009).

II. Econometric Model

To statistically test for gender differences in absence patterns, IM estimate the Cox proportional hazard model shown in Equation 1, where t indexes days from the start of the previous absence spell, X_{it} are covariates, and $\mathbf{\Psi}$ is a vector of coefficients.[4] There are three main control variables: (1) an indicator for whether the teacher is female (F_i), (2) an interaction of female with an indicator for a distance of 28 days (M_{it}), and (3) an interaction of the female indicator with an indicator for distances that are multiples of seven (S_{it}). Z_{it} is a vector of controls for day of the week and worker characteristics.[5]

(1) $$h(t, X_{it}, \Psi) = \lambda(t) exp\left(\alpha + \beta F_{it} + \gamma M_{it} F_i + \delta S_{it} F_i + +\theta Z_{it}\right)$$

The main parameter of interest is γ, which measures the difference in the absence hazard rates of women and men 28 days after the start of a previous spell, after allowing for both a different baseline hazard (β) and seven-day periodicity (δ) for women. IM's inclusion of a

[4] There are advantages and disadvantages to this model. One advantage is the simple way in which the model controls for the persistence of absences; absence spells are treated as left truncated (in other words, not in the risk set for a new absence spell) until the worker returns to work, and there is a flexibly estimated baseline hazard for each interval of time. Using the distance between the starts of absence spells also conceptually corresponds to the distance between menstrual cycles, measured as the number of days between their onsets. An alternate approach is a probit model that estimates the probability of an absence given a previous absence 28 days ago. This model would capture more than just the distances between the starts of consecutive spells but requires some alternate means to handle persistence (for example, including many lags of previous absences). We do not have a strong argument for preferring this approach since it also has significant drawbacks (for example, insufficient power to precisely estimate many lags).

[5] In the Italian bank data, worker characteristics include: age, years of schooling, marital status, number of children, managerial occupation, and seniority. In the NYC teacher data, worker characteristics include: age, teaching experience, and education (masters degree).

separate seven-day periodicity term for women is based on their observation from the graphical analyses that there are large spikes in the Kaplan-Meier hazard rates for both genders at distances that are multiples of seven. Since they attribute these spikes to the high frequency of absences on Mondays (the "Monday morning" effect) and family and non-work commitments that repeatedly fall on the same day of the week, they reason that the observed seven-day periodicities are constant but gender-specific (for example, women being more likely to have family commitments that fall on the same day of the week, like driving their children to soccer practice).

In contrast, we believe the spikes at seven-day intervals are primarily mechanical effects of the five-day workweek. Like schools and many other firms, most Italian banks do not open on weekends, though some open for a shortened business day on Saturday. In the Italian bank data, only 0.26 percent of absence spells begin on Saturday or Sunday. Thus, conditional on the date of a prior absence, the probability that the bank is open seven days later (or any multiple of seven days later) is considerably higher than for distances that are not multiples of seven. A key distinction between our explanation and IM's is that we do not assume that the differences between the seven-day periodicities generated by the five-day workweek are constant.[6]

Figure 1 uses simulated absence data to illustrate how the five-day work week could generate non-constant, seven-day periodicities; it plots the Kaplan-Meier hazard rates and their differences for two data sets of 10,000 males and 10,000 females over 1,000 days where absence probabilities are set at about five percent for men and seven percent for women. In the

[6] It is worth noting that Schempp (2009) makes a similar point regarding the importance of the 5-day workweek in his interpretation of the patterns found in the Italian bank data, though he does not consider its implications for estimated gender differences in hazard rates at multiples of 7 days.

simulations, the health of worker i of gender g on day t (H_{igt}) follows an AR-1 process with an independent and normally distributed error term with mean zero and variance one:

(2) $H_{igt} = \rho_g H_{ig,t-1} + \varepsilon_{igt}, \quad \varepsilon_{igt} \sim N(0,1)$

Workers are sick if their health falls below a sickness threshold (α_g), and once sick, do not recover until their health exceeds a separate wellness threshold (λ_g), where $\alpha_g \leq \lambda_g$.[7] Absences are only observed when workers are sick on workdays, and the rates of health shock persistence (ρ_g), and the thresholds, α_g and λ_g, may differ by gender.

The left panel of Figure 1 displays a simulation where there is no persistence in health (ρ_g = 0) and the sickness and wellness thresholds are the same for each gender ($\alpha_g = \lambda_g$), while the right panel displays a simulation where these parameters differ.[8] Notably, both simulations show large spikes at multiples of seven, despite our lack of assumptions about gender differences in the absence probabilities on particular days of the week or at these distances.

More importantly, while the gender difference in the periodicity of absence spells at multiples of seven appears constant when health is not persistent (left panel), it is clearly not constant when health is persistent and recovery rates differ (right panel).[9] This motivates two concerns: First, IM's assumption of a constant seven-day periodicity may be incorrect and their

---

[7] The separate sickness and wellness thresholds replicate the empirical fact that workers usually wait between initiating absence spells.

[8] In the first simulation, for men, $\rho_m = 0$ and $\alpha_m = \lambda_m = -1.645$, and for women, $\rho_f = 0$ and $\alpha_f = \lambda_f = -1.476$. In the second simulation, $\rho_m = 0.78$, $\alpha_m = -3.1$, $\lambda_m = -1.6$, and $\rho_f = 0.75$, $\alpha_f = -2.8$, $\lambda_f = -1.0$.

[9] To mitigate the weekend effect, one could convert the data to a 5-day week (for instance, Monday to Monday is a 5 day interval) and look for a 20-day instead of 28-day spike; however, we demonstrate in the Appendix, this will not necessarily eliminate the spikes induced by five-day workweek.

model may therefore be mis-specified. Second, and more worrisome, differences in the absence hazard rates at 28 (a multiple of seven) may be attributable to factors other than menstruation (such as differences in persistence or recovery rates). The first concern can be addressed by testing the joint restriction that the coefficients on the separate interactions between female and each multiple of seven, excluding 28, are equal. We evaluate the empirical importance of the second concern by comparing the absence hazard rates of two groups who do not experience menstruation: older and younger men.

III. Re-analysis of Absenteeism in an Italian Bank

Ichino and Moretti analyze data covering the absences of employees in a large Italian bank over a period of three years, and we refer the reader to their paper for additional details on these data. Using data and code generously provided to us by Ichino and Moretti, we replicate the results of the hazard regressions using IM's most rigorous specification, which includes controls for female, an interaction of female with a distance of 28 days, an interaction of female with a distance that is a multiple of seven, workers' characteristics (age, years of schooling, marital status, number of children, managerial occupation, and seniority) and day of the week.

In Column 1 of Table 1, we display IM's published estimates for workers under the age of 45. All estimates are reported as hazard ratios, with coefficients greater (less) than one indicating a positive (negative) effect, and t-statistics are reported in the parentheses. In line with females having a higher number of absence spells, the coefficient on female is significantly greater than one. More importantly, the coefficient on the interaction of female with a distance of 28 days is 1.15 and statistically significant with a t-statistic of 2.16. This is the main piece of evidence supporting IM's conclusion that, relative to younger men, younger women are more likely to start absence spells 28-days apart.

However, there were several anomalies in the computer code used to estimate this regression. First, there are some adjacent absence spells that are separated by a distance of one, which is inconsistent with the coding of consecutive days of absence into spells.[10] Second, when workers had multiple absence spells, their last absence spell was unintentionally dropped from the regression, rather than being coded as right censored.[11] Third, controls for day of the week on day *t* in the hazard regression were coded to control for the day of the week on which the previous absence spell started.[12] Finally, employees were coded as having their age at the start of the three-year sample period, not their actual age.

We present hazard regressions that correct all these coding errors – one-day adjacent spells, right censoring, day of the week controls, and age – in Column 2. Correcting these errors decreases the coefficient on the interaction of female and 28 days to 1.11 and causes its t-statistic to fall to 1.51. Surprisingly, the changes in the age coding, rather than the day of the week controls, account for most of this reduction; we believe this is due to slight changes in the

[10] The distance between spells should never equal one since consecutive days absent are counted as a single spell. Any adjacent spell that is separated by a distance of one from a previous spell should have been part of the previous spell.

[11] This is true for all workers with multiple absences. For workers with a single absence spell, their spell was correctly treated as right censored.

[12] Suppose an absence spell started on a Monday and there were 30 days until the start of the next spell. IM's code would create 30 observations with hazard time running from 1 to 30, but all of the observations would be coded as Mondays. We fix this, so that the observation at time = 2 occurs on a Tuesday, time = 3 occurs on a Wednesday, etc. This is potentially important because the hazard rate for an absence on the weekend is nearly zero.

composition of the under 45 sample. In these data, men who turned 45 in 1993 or 1994 are considerably more likely to have absences at distances of 7, 14, and 21 days; including them in the under 45 sample causes the negative interaction of female and multiple of seven to be larger in magnitude and thus the positive interaction of female and 28 is larger in magnitude as well.[13]

Another issue with these hazard regressions is that while the distances between the start of absence spells are likely correlated within individuals, IM treat all spells of absence as independent. Although this is not an error in coding, we believe a more appropriate treatment of the data is to calculate standard errors allowing for clustering at the individual level. When we do so, the t-statistic on the interaction of female and 28 days falls to 1.39 (Table 1, Column 3).

Next, we turn to model specification issues; recall that the inclusion of an interaction of female with the distance being a multiple of seven days was motivated by the potentially incorrect assumption of a constant, gender-specific seven-day periodicity. In practice, including this interaction term could bias IM towards finding a positive effect on female and 28 days if the relative female hazard rates are very low at early multiples of seven. The coefficient on the interaction of female with 28 days is identified as the difference between the relative female

[13] Since correction of the age variable alters the sample of employees under age 45, we test the sensitivity of the results to sample composition by estimating hazard regressions using age cutoffs from 42 through 52; Appendix Table 1 shows estimates using the regression specification from Column 3 of Table 1. The coefficient on the interaction of female and 28 days is only statistically significant at conventional levels when we use age cutoffs of 50 or 51. If we use the specification in Column 4 of Table 1, the interaction of female and 28 days is never significant, regardless of the age cutoff.

hazard rate at 28 and that at other multiples of seven, and the latter is weighted towards early multiples of seven since observations are decreasing in *t*.

If we replace the interaction between female and *any* multiple of seven days with separate interactions between female and *each* multiple of seven (Table 1, Column 4) the coefficient on female and 28 days falls to 1.09 and its t-statistic to 1.21. A Wald test rejects the equality of the interactions of female and multiples of seven, excluding 28, with a p-value below 0.0001.[14] This confirms that the gender difference in seven-day periodicity is not constant in these data, and reveals positive interactions between female and 63 and 70 days which are more precisely estimated and of a greater magnitude than the coefficient on the interaction between female and 28 days, even though menstrual cycles are extremely unlikely to occur at 63 or 70 day intervals.[15]

Results using the same specification for men and women aged 45 and over are shown in Column 6, alongside IM's published results for this age group in Column 5. Like IM, we find

[14] Ichino and Moretti include an indicator for any multiple of 7 than or equal to 70, so we include interactions with each multiple of 7 up to 70. Replacing the interaction between female and multiple of 7 days with these separate interactions in IM's original code (namely, without correcting the coding errors or clustering the standard errors) causes the coefficient on female and 28 days to drop to 1.10 with a t-statistic of 1.47.

[15] We gauge the likelihood of menstrual cycles occurring at distances of 63 and 70 days using information from Chiazze et al. (1968), who collected data on onset of each menstrual cycle between January 1964 and December 1965 for a group of 2,316 (predominantly Catholic) American and Canadian women, each of whom were followed for a minimum of ten cycles. We use the probability distribution of cycle lengths reported in their study to estimate the probability that a woman has two (possibly non-consecutive) menstrual cycles at various distances. We estimate that the probability of cycles at distances of 28 days and 56 days are, respectively, 16 and 9 percent. In contrast, estimated probabilities of cycles at distances of 63 and 70 days are just 2.6 and 0.8 percent, respectively.

that older women are no more likely than older men to have absence spells initiating 28-days apart, but that older men are more likely to have absence spells initiating 14-days apart.

The large gender differences in hazard rates between men and women at multiples of seven other than 28 suggests caution in attributing differences at 28 days to menstruation, but it is possible that menstruation-related absences could affect the entire survival curve. To gauge the importance of confounding seven-day periodicities while abstracting away from menstruation, we estimate hazard regressions for two groups who do not experience menstruation: men under 45 and men 45 and older. For these specifications, we include (1) an indicator variable for being age 45 and over, (2) interactions between this indicator variable and distances that are multiples of seven, and (3) the controls used in the previous hazard regressions (for example age, experience, day of the week, etc.).[16]

We present these results in Column 7. Nearly all of the interactions between the indicator for older males and multiples of seven are statistically significant, signifying the potential importance of the distortions induced by the five-day workweek. Moreover, the coefficient on the interaction between older males and 28 days is positive, statistically significant, and at least as large as the coefficient for women in Column 4. While the interaction between older males and 28 days is not the largest of all the interaction terms in Column 7, we can think of no explanations for why older male employees in an Italian bank are much more likely than younger males to have absence spells initiating in 7- or 14-day cycles.

[16] These regressions compare two different age ranges, rather than two genders of the same age range. We therefore also estimated specifications that allow for a separate age trend for males 45 and older and find very similar results to those reported here; when this additional age control is included the coefficient on older male decreases but the coefficients on the interactions between multiples of 7 and older male remain quite stable.

The results presented in this section do not support the finding of a link between menstruation and female absenteeism at work in the Italian bank data. The main results and conclusions presented in earlier work do not hold up to corrections of coding errors and modest sensible modifications in regression specification. Using simulations as well as “placebo” comparisons of younger and older males, we also demonstrate that gender differences in hazard rates at distances that are multiples of seven days may be spurious effects of the five-day workweek.

IV. Absences Among New York City Teachers

The sensitivity of the main results from IM’s analysis of the Italian bank data, as well as and the presence of significant differences between the hazard of absences spells among younger and older males at various multiples of seven casts doubt on the interpretation of 28-day cycles of absence spells among females as due to menstruation. Nevertheless, the estimated interaction of female with a distance of 28 days is still positive and one might interpret this as providing some support for the link between menstruation and absenteeism. To see if this result holds in a different dataset with a much larger sample size, we conduct a similar analysis of absences among New York City teachers.[17]

New York is the largest school district in the United States and employs roughly 80,000 teachers annually to staff 1,500 schools. All teachers are employed full-time under the same collectively bargained contract and are paid based on a salary schedule that depends only on their years of experience and graduate education. Thus, there is no gender wage gap for teachers,

[17] While it would be preferable to examine a wide range of professions, daily data on absences for a large sample of individuals are rare. Thus, like IM, we take advantage of data that we have used for research on another topic.

conditional on experience and education. Teachers earn ten days of paid absence for each year of work. Unused days roll over and teachers can accumulate up to 200 days of absence, but teachers cannot take more than ten undocumented days of paid absence per school year. If a doctor's note certifies that an absence was due to illness, then the absence will not count towards the annual cap. Teachers who resign or retire are paid 1/400th of their salary for each of their accumulated unused absence days, so using "paid" absences entails a future financial cost.

We use data on all absences taken by all full-time public school teachers in New York City during the school years 1999-2000 through 2002-2003, as well as teacher characteristics (demographics, education, and experience) and extended leaves taken during this time period (such as sabbatical or maternity leave). Kane, Rockoff, and Staiger (2008) and Herrmann and Rockoff (2010) provide detailed descriptions of these data.

We can distinguish absences taken for a number of special reasons, such as jury duty, military service, funeral, or religious holiday. These comprise 24 percent of absences and, to be more in line with the analysis of Italian bank data studied in Section 2, we remove them from the analysis and focus on absences taken for medically certified and uncertified illness and personal business.

A. Summary Statistics

Summary statistics for the 67,719 female teachers and 25,789 male teachers used in our analysis sample are shown in Table 2.[18] Following IM, we excluded "all employees who took maternity leave at any point," as well as any teacher who took an extended leave of absence (medical, sabbatical) during the school year or left their teaching position before the end of the school year. We also excluded teachers with zero absences, since we can only conduct the

[18] Summary statistics for the full sample of teachers are available from upon request.

analysis for teachers with at least one absence spell during the school year. As in IM, we separately examine younger (under 45) and older (45 and over) workers.

Teaching has historically been one of the most common female professions, and it is not surprising that we see a clear majority of women among both older and younger teachers (71 and 74 percent, respectively). Within age categories, teachers of both genders are similar in their average age and years of teaching experience, and females tend to be somewhat more likely to have a master's degree. Rates of absence are higher for younger women than young men (7.84 vs. 7.20 per year) and higher for older women than older men (7.91 vs. 7.82 per year). Thus, the stylized fact that women are absent more often than their male colleagues holds for our sample of public school teachers, although the gaps are wider in the Italian bank examined by IM.[19] Consecutive days of absence are grouped into spells in our analysis, and we present the average annual number of spells of absence by gender and age category. While younger women have more individual days of absence than younger men, they have slightly fewer spells per year (5.65 vs. 5.76). Thus, compared to younger men, younger women have fewer but longer absence spells; 16 percent of younger women's spells last more than one day, compared to 13 percent of younger men's. This is also true for older women, who have slightly fewer spells than older men per year (5.44 vs. 5.55), and their spells are longer; 19 percent of older women's absence spells last more than one day, compared to 16 percent of those of older men.[20]

---

[19] In their data, female employees averaged roughly 13 absences per year compared to only 8 for males.

[20] Absences and absence spells are more likely to occur on certain days of the week; for absences, 21 percent occur on Monday, 20 percent on Tuesday, 18 percent on Wednesday, 17 percent on Thursday, and 24 percent on Friday. For absence spells, 24 percent start on Monday, 19 percent start on Tuesday, 17 percent start on Wednesday, 17 percent start on Thursday, and 23 percent start on Friday.

B. Hazard Regressions

Our regression analysis of absence spells among teachers closely follows our previous analysis of the Italian bank data. In Table 3, we present the results from specifications that include either an indicator for female or an indicator for being age 45 and older, separate interactions of this indicator and each multiple of seven, and controls for day of the week and teacher characteristics. The first two columns report the within age group comparisons of female and male teachers, while the last two columns report the within gender comparisons of teachers under 45 and those 45 and older. For all groups, Wald tests strongly reject the equality of the coefficients on the interactions of the group indicator (female or age 45 and older) with distances that are multiples of seven other than 28. The p-values of these tests are all less than 0.0001, suggesting that empirically, group differences in seven-day periodicity are seldom constant. While we do not push a causal interpretation of these estimates, we find that most of the interactions between the group indicator and multiples of seven are statistically significant, and if anything, male teachers are *more* likely than their female colleagues to initiate absence spells in 28-day cycles.

An important feature of this analysis is our large sample size, so the lack of any evidence of a menstruation-absence link among New York City teachers is not due to lack of power. While the econometric difficulties associated with the five-day work week remain, the results presented above, in conjunction with the re-analysis of the Italian bank data, suggest that there exists no solid empirical evidence in favor of the biological-based explanation for increased work absenteeism among women.

V. Conclusion

A link between menstruation and workplace absenteeism among females provides a provocative and potentially important biological explanation for gender gaps in labor market outcomes. However, this explanation lacks empirical support. We find no evidence that females are more likely to initiate absence spells at 28-day intervals in either our re-analysis of the Italian bank data used by Ichino and Moretti or in a large dataset on public school teachers in New York City. Moreover, the fact that the five-day workweek accentuates differences in absence hazards between groups at multiples of seven suggests that gender differences in the hazard rate of absences spells at 28 days cannot necessarily be interpreted as an effect of menstruation.

Because of these econometric difficulties, we doubt that effects of menstruation will be accurately measured only with data on absences. A more promising approach is to collect information on both absences from work and menstruation (or menstruation-related symptoms). There are many studies in the medical literature that measure work absences due to (pre-) menstrual symptoms using survey data, and three to 16 percent of employed women in these (often non-representative) samples report having missed work in the previous year due to menstruation.[21] While this literature provides a rough sense of the fraction of women ever missing work due to menstruation, we have little evidence on the fraction of work days missed.

[21] See Halbriech et al. (2003) for a review. Interestingly, Ichino and Moretti cite a paper in this literature (Chawla et al. 2002) as providing evidence in line with their estimated number of menstrual cycle-related absences. However, IM focus on the paper's finding that women with severe symptoms experienced additional "cut down days," defined as "days that were cut-down on usual activities for a half-day or more" and which are explicitly distinguished from lost time at work. In fact, Chawla et al. (2002) find "little evidence that women spent time in bed or reduced their time at work, or activities at home or school caused by premenstrual symptoms or PMDD." PMDD stands for premenstrual dysphoric disorder, a diagnosis associated with severe premenstrual symptoms estimated to affect 3-8 percent of the female population between ages 18 and 48 (Katz et al. 2007).

This is precisely the point made by Oster and Thornton (2009, 2011) with regard to menstruation and girls' absences from school. In that literature, survey evidence had been used to support the idea that availability of sanitary products was a major impediment to girls' attendance (World Bank 2005). Indeed, Oster and Thornton also find that a high fraction of the girls they study in Nepal report having missed school due to menstruation (35 percent). However, using daily data on menstrual cycles and absenteeism, they find that menstruation can account for just 1.5 percent of girls' annual school absences.

While generalization to other contexts is unclear, several factors suggest that Oster and Thornton's estimate of a 1.5 percent increase in absence due to menstruation could overstate the true effect of menstruation on women's absences from work in developed countries. For example, they focus on girls in a developing country with limited access to sanitation products or medical technology, some of whom have only recently reached menarche. Yet the true impact of menstruation of work absence would have to be an order of magnitude greater in order to explain gender gaps in absenteeism in developed countries, where absence rates for females are typically found to be at least 25 percent greater than for males (see, for example, Leigh 1983, Vistnes 1997, Mastakaasa and Olsen 1998, Bridges and Mumford 2001, or Broström Johansson, and Palme 2002).

Though existing evidence does not support a large impact of menstruation on female work absences, we still believe that more research on this topic is needed from different countries and occupations. For example, labor laws and labor contracts which recognize the right of women to take a "feminine day" or "menstrual leave" once per month are common in Indonesia, Japan, South Korea, and Taiwan. These labor institutions may be important in mediating the relationship between menstruation and labor market outcomes among women.

References

Altonji, Joseph, and Rebecca Blank. 1999. "Gender and Race in the Labor Market." In *Handbook of Labor Economics.* Vol. 3C, ed. Orley C. Ashenfelter and David Card, 3143-3259. New York: Elsevier Science.

Blau, Francine, and Lawrence M. Kahn. 2000. "Gender Differences in Pay," *Journal of Economic Perspectives* 14(4): 75-99.

Bridges, Sarah, and Karen Mumford. 2001. "Absenteeism in the UK: A Comparison Across Genders," *The Manchester School*, 69(3): 276–284.

Broström, Göran, Per Johansson, and Mårten Palme. 2002. "Economic Incentives and Gender Differences in Work Absence Behavior," IFAU Working Paper 2002:14.

Chawla, Anita, Ralph Swindle, Stacey Long, Sean Kennedy, and Barbara Sternfeld. 2002. "Premenstrual Dysphoric Disorder: Is There an Economic Burden of Illness?" *Medical Care*, 40(11): 1101-1112.

Chiazze, Leonard, Franklin T. Brayer, John Macisco., Margaret Parker, and Benedict Duffy. 1968. "The Length and Variability of the Human Menstrual Cycle," *Journal of the American Medical Association*. 203(6): 377-380.

Goldin, Claudia. 1990. "Understanding the Gender Gap: An Economic History of American Women," in Equal Employment Opportunity, Labor Market Discrimination and Policy Paul Burstein ed. pages 17-26.

Halbriech, Uriel, Jeff Borenstein, Terry Pearlstein, and Linda S. Kahn. 2003. "The Prevalence, Impairment, Impact and Burden of Premenstrual Dysphoric Disorder (PMS/PMDD)," *Psychoneuroendocrinology*, 28(3): 1-23.

Herrmann, Mariesa, and Jonah Rockoff. 2010. "Worker Absence and Productivity: Evidence from Teaching," NBER Working Paper.

Ichino, Andrea. and Enrico Moretti. 2009. "Biological Gender Differences, Absenteeism and the Earnings Gap," *American Economic Journal: Applied Economics* 1(1): 183-218.

Kane, Thomas, Jonah Rockoff, and Douglas Staiger. 2008. "What Does Certification Tell us about Teacher Effectiveness? Evidence from New York City," *Economics of Education Review*, 27(6): 615-631.

Katz, Vern, Gretchen Lentz, Rogerio Lobo, and David Gershenson. 2007. Comprehensive Gynecology, 5th Edition. Philadelphia: Mosby, Elsevier.

Leigh, J. Paul. 1983. "Sex Differences in Absenteeism," *Industrial Relations*, 22(3): 349-361.

Mastekaasa, Arne, and Karen Modesta Olsen. 1998. “Gender, Absenteeism, and Job Characteristics: A Fixed Effects Approach,” *Work and Occupations*, 25(2): 195-228.

Oster, E. and Thornton, R. (2009) “Menstruation and Education in Nepal,” NBER Working Paper.

Oster, Emily, and Rebecca Thornton. 2011. “Menstruation, Sanitary Products and School Attendance: Evidence from a Randomized Evaluation,” *American Economic Journal: Applied Economics*, 3(1): 91-100.

Paringer, Lynn. 1983. “Women and Absenteeism: Health or Economics?” *American Economic Review* Papers and Proceedings, 73(2): 123-127.

Schempp, Paul (2009) “Defending Italian Bank Tellers,” Unpublished Manuscript, University of California, Berkeley.

Vistnes, Jessica Primoff. 1997. “Gender Differences in Days Lost from Work Due to Illness,” *Industrial and Labor Relations Review*, 50(2): 304-323.

Table 1
Hazard of Absence Spells - Italian Bank Data

Panel A:	Under 45				45 and Over		*Panel B:*	Men	Women
Female vs. Male	IM (2009)	Re-analysis			IM (2009)	Re-analysis	*Older vs Younger*	Re-analysis	Re-analysis
	(1)	(2)	(3)	(4)	(5)	(6)		(7)	(8)
Female	1.39	1.47	1.47	1.47	1.29	1.33	Older	0.93	0.84
	(35.94)	(41.64)	(20.38)	(20.38)	(13.82)	(7.43)		(-2.66)	(-3.62)
Female*28 Days	1.15	1.11	1.11	1.09	0.99	1.02	**Older*28 Days**	1.16	1.04
	(2.16)	(1.51)	(1.39)	(1.21)	(-0.08)	(0.15)		(2.06)	(0.32)
Female*Multiple of 7 Days	0.95	0.98	0.98		0.98				
	(-2.04)	(-0.70)	(-0.63)		(-0.43)				
Female*7 Days				0.78		0.85	Older*7 Days	1.38	1.42
				(-2.71)		(-1.07)		(3.51)	(2.27)
Female*14 Days				0.94		0.72	Older*14 Days	1.38	1.01
				(-0.88)		(-2.81)		(4.64)	(0.10)
Female*21 Days				0.84		1.01	Older*21 Days	1.12	1.29
				(-2.60)		(0.09)		(1.77)	(2.19)
Female*35 Days				1.00		0.90	Older*35 Days	1.11	0.96
				(0.00)		(-0.80)		(1.55)	(-0.30)
Female*42 Days				1.05		1.10	Older*42 Days	0.96	0.96
				(0.70)		(0.74)		(-0.57)	(-0.31)
Female*49 Days				1.10		1.00	Older*49 Days	1.07	0.94
				(1.17)		(-0.02)		(0.90)	(-0.44)
Female*56 Days				1.08		1.28	Older*56 Days	0.99	1.14
				(0.89)		(1.58)		(-0.10)	(0.83)

Female*63 Days			1.18	1.13	Older*63 Days	1.26	1.16
			(1.92)	(0.86)		(2.68)	(1.04)
Female*70 Days			1.23	1.08	Older*70 Days	1.05	0.89
			(2.23)	(0.41)		(0.55)	(-0.60)
Correction for Miscoding	√	√	√	√		√	√
Clustered Standard Errors		√	√	√		√	√
Chi-squared Statistic on Wald test			31.95	15.88		21.31	11.80
P-value on Wald test			0.0001	0.0442		0.0064	0.1605

Note: The first and fifth columns display results from Ichino and Moretti (2009), Table 2. Corrections for miscoding are related to one-day adjacent spells, right censoring, day of the week controls, and age. Specifications that allow for clustering of residuals do so at the the individual worker level. All specifications control for age, years of schooling, marital status, number of children, managerial occupation, seniority, and dummies for each day of the week. The Wald test tests the joint restrictions that the coefficients on all the interactions of female (older) with distances that are multiples of 7, excluding 28, are equal. T-statistics are shown in parentheses. A hazard ratio of 1 indicates no effect.

Table 2
Summary Statistics on New York City Teachers by Gender and Age Group

	Under Age 45		Age 45 and Over	
	Male	Female	Male	Female
Number of Teachers	14,538	40,790	11,251	26,929
Number of Observations (Teacher-Year)	34,801	100,498	33,339	87,455
Age	33.6	32.7	53.1	52.9
Teaching Experience	4.3	4.4	14.7	13.7
Black	22.8%	24.0%	15.7%	20.4%
Hispanic	16.7%	17.8%	8.7%	10.3%
Masters Degree	35.1%	40.4%	23.4%	27.9%
Number of Days Absent	7.20	7.84	7.82	7.91
Number of Absence Spells	5.76	5.65	5.55	5.44
Number of 1-Day Absence Spells	5.02	4.77	4.68	4.42
Number of 2-3-Day Absence Spells	0.62	0.70	0.69	0.80
Number of 3+ Day Absence Spells	0.12	0.18	0.18	0.22
Percent with only 1 Absence Spell (Right censored)	7.3%	5.8%	8.6%	6.9%

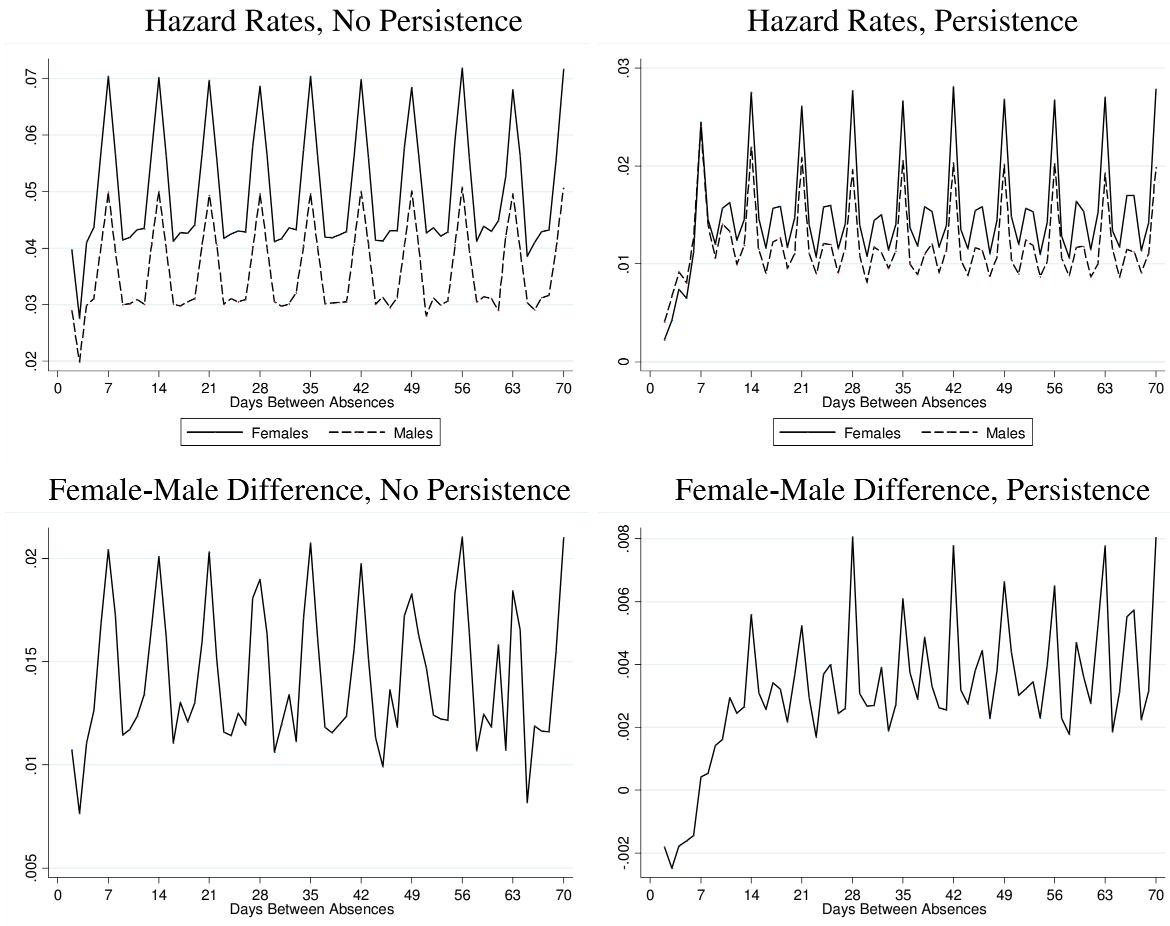
Note: The unit of observation for the calculataion of average characterisitcs is a teacher-year.

Absences only include those for illness (certifed by a doctor's note or not) and personal reasons.

Table 3
Hazard of Absence Spells - NYC Teachers

Panel A: Males vs. Females	Under 45 (1)	45 and Over (2)	*Panel B: Older vs. Younger*	Men (3)	Women (4)
Female	0.99	0.99	Older	0.92	0.88
	(-0.98)	(-2.34)		(-7.01)	(-17.28)
Female*28 Days	0.90	0.89	**Older*28 Days**	1.03	1.03
	(-6.18)	(-6.28)		(1.39)	(1.89)
Female*7 Days	0.92	0.84	Older*7 Days	1.38	1.27
	(-4.42)	(-9.25)		(13.79)	(16.55)
Female*14 Days	0.89	0.87	Older*14 Days	1.10	1.09
	(-6.80)	(-7.48)		(4.48)	(6.28)
Female*21 Days	0.90	0.89	Older*21 Days	1.05	1.05
	(-6.41)	(-6.57)		(2.38)	(3.50)
Female*35 Days	0.90	0.90	Older*35 Days	0.95	0.96
	(-5.90)	(-5.06)		(-2.18)	(-2.72)
Female*42 Days	0.92	0.96	Older*42 Days	0.90	0.94
	(-3.98)	(-1.72)		(-4.07)	(-3.66)
Female*49 Days	1.02	0.98	Older*49 Days	0.93	0.90
	(0.80)	(-0.83)		(-2.52)	(-6.12)
Female*56 Days	0.97	0.96	Older*56 Days	0.92	0.91
	(-1.00)	(-1.38)		(-2.50)	(-4.48)
Female*63 Days	1.01	1.03	Older*63 Days	0.95	0.97
	(0.26)	(0.84)		(-1.47)	(-1.38)
Female*70 Days	0.98	1.04	Older*70 Days	0.90	0.96
	(-0.59)	(1.00)		(-2.30)	(-1.41)
Chi-squared Statistic on Wald Test	41.90	64.87		228.32	362.96
P-value on Wald test	0.0000	0.0000		0.0000	0.0000

Note: All specifications control for age, teaching experience, education (masters degree), and indicator variables for day of the week. The Wald test tests the joint restrictions that the coefficients on all the interactions of female (older) with distances that are multiples of 7, excluding 28, are equal. Standard errors allow for clustering at the individual worker level. T-ratios are shown in parentheses. A hazard ratio of 1 indicates no effect.



Note: Simulated data sets are for 10,000 males and 10,000 females over 1,000 days where absence probabilities are set so that men are absent roughly 5 percent of days and women are absent roughly 7 percent of days. Health follows an AR-1 process with a normally distributed i.i.d error term: $H_{igt} = \rho_g H_{igt-1} + \varepsilon_{igt}$, $\varepsilon_{igt} \sim N(0,1)$. Workers are sick if their health falls below a threshold (α_g), and once sick, do not return to work until their health exceeds a separate threshold (λ_g), where $\alpha_g \leq \lambda_g$. Absences are only observed when workers are sick on workdays, and the rates of health persistence (ρ_g) and thresholds, α_g and λ_g, may differ by gender. In the "No Persistence" simulations, $\rho_m = \rho_f = 0$, $\alpha_m = \lambda_m = -1.645$, and $\alpha_f = \lambda_f = -1.476$. In the "Persistence" simulations, $\rho_m = 0.78$, $\alpha_m = -3.1$, $\lambda_m = -1.6$, $\rho_f = 0.75$, $\alpha_f = -2.8$, $\lambda_f = -1.0$.

Figure 1
Kaplan-Meier Hazard Rates for Simulated Absence Data