Biological Gender Differences, Absenteeism, and the Earnings Gap: Comment

Jonah Rockoff
Columbia Business School

Mariesa Herrmann
Columbia University

December 2009*

Abstract

Ichino and Moretti (2009) find that much of the gender gap in absenteeism at an Italian bank is explained by absences with a 28-day cycle. This is interpreted as an effect of menstruation which subsequently explains part of the gender earnings gap. We find their results are not robust to the correction of program errors and allowing for serial correlation. We also find that differences between pre-menopausal women and same-aged men in absences cycles around 28 days are smaller than differences between older and younger men. We conclude there is little evidence that menstruation explains gender gaps in absenteeism and earnings.

---

* Correspondence should be sent to jonah.rockoff@columbia.edu. We thank Andrea Ichino and Enrico Moretti for sharing their data and computer code and for helpful discussions. We 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 Hermann's work is supported by a National Science Foundation Graduate Research Fellowship. Any opinions, findings, conclusions or recommendations expressed in this study are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. All errors are our own.

A large literature in economics has documented differences in earnings between men and women (see Goldin (1990), Blau and Kahn (2000)). Standard explanations for these differences include gender differences in preferences, gender differences in skills, and discrimination (Altonji and Blank (1999)). In a recent paper, Ichino and Moretti (2009), hereafter IM, provide a biological explanation for part of the gender earnings gap: female absenteeism caused by the menstrual cycle. Their evidence in favor of this explanation comes from comparing the absence patterns of male and female employees of a large Italian bank. IM show that the hazard rate of a new absence spell increases significantly for females under age 45 (relative to males under 45) 28 days after the beginning of a previous absence spell, and that this pattern does not exist for employees 45 years and older, when women are likely to be approaching menopause.[1] Their estimates suggest that the "additional absenteeism induced by the menstrual cycle" explains one-third of the gender gap in days of absence and 14 percent of the gender gap in earnings.

We find that the key results in IM are not robust to the correction of errors in their computer programs and calculating standard errors clustered by individual. Pre-menopausal female Bank employees are not significantly more likely than same-aged male colleagues to have absences 28 days apart. We find weak evidence that younger women have more absences than younger men in cycles *around* 28 days, but this pattern arises even more strongly in a comparison of older men with younger men. These results caution against concluding that menstruation plays an important role in explaining gender differences in labor market outcomes.

2. Do Female Absences Follow the Menstrual Cycle Among Italian Bank Employees?

In this section, we revisit the main pieces of graphical and statistical evidence from IM that suggest the menstrual cycle is an important cause of work absence by female Italian bank

[1] A spell is a continuous period of absence.

employees. In Figure 1 of their paper, IM provide graphical evidence that women are more likely than men to be absent 28 days apart. Specifically, the figure shows a spike at 28 days in the gender difference in the probability distributions of distance between consecutive spells of absence. We reproduce this result (Figure 1) and also plot the male and female probability distributions of absences.[2]

In addition to the spike at day 28, the figure also shows that the densities of the distributions for both men and women are unusually high for distances that are multiples of seven. This is mostly a mechanical effect of the five-day workweek.[3] Any day that is a distance of seven (or a multiple of seven) days from a workday must fall between Monday and Friday, but distances not divisible by seven can fall on a Saturday or Sunday. To illustrate this more clearly, we simulate data for workers absent on random dates; females are absent on work days with a 5 percent probability and males are absent on work days with a 3 percent probability. Density plots for these data (Figure 2) also show peaks on seven day intervals.

Importantly, differences in the probability distributions between males and females will be greater at distances with greater probability mass. For example, even if the true gender difference in the probability of absence were the same on days 27 through 29, the difference in the probability density would be greater on day 28 than on the adjacent days. Another potentially important issue with this graphical analysis is that it does not take account for the fact that a worker cannot begin a new absence spell until one day after their previous spell is complete. Figure 3 presents the probability densities (and gender differences) by distance from

[2] Our figure is slightly altered from Figure 1 in IM. We do not rescale the male-female difference to have a mean of 0 between 0 and 50 days, and we combine spells that form a continuous period of absence, so the distance between spells can never equal one day.

[3] To the best of our knowledge, most Italian banks are not open on weekends, though some open for a shortened business day on Saturday. This is supported by the fact that, in the data used by IM, only 0.19 percent of absence spells for Italian bank employees start on weekends. IM attribute the spiking at 7-day multiples to people taking absences on the same day of the week (e.g., a “Monday effect”). While this behavior will also increase density at 7-day multiples, in practice this is unimportant relative to the five-day workweek.

the previous spell after accounting for weekends and the length of the prior spell.[4] Here we see no spike in the gender difference at a distance of 28 days.

Proportional hazard regressions should property account for the issues of weekends and previous spells. The results in the regressions presented in IM indicate that the hazard rate of an absence spell increases 28 days after the start of a previous absence spell for females, relative to males, under the age of 45. Using their data and code, we replicate this result in Column 1 of Table 1. This specification includes controls for a worker's characteristics (age, years of schooling, marital status, number of children, managerial occupation, and seniority) and indicators for day of the week.

There were three errors in the computer code used by IM to estimate these regressions. First, employees were coded as always being the same age as they were at the start of the sample period, as opposed to their actual age. Thus, age does not vary within employees over time, and some employees who were well over 45 during the latter part of the sample were always included in the "under 45" group. Second, all days between adjacent spells were coded as falling on the day of the week corresponding to the start of the previous spell, so controls for day of the week did not accurately capture the changes in the probability of an absence between, say, Mondays and Saturdays. Third, the absence spell of every worker with only one absence spell were coded as right censored, but the last absence spell of every other worker, which is also right censored, was simply dropped from the regression.

We estimate hazard regressions that correct these coding errors but use the same set of control variables.[5] Additionally, because the length of time between absence spells is highly

[4] Specifically, we define the density at a given distance t as the number of new absence spells occurring t days from the start of a prior absence spell divided by the total number of spells in the data set for which t days from the start of the prior absence spell did not fall on a weekend and for which the prior absence spell lasted less than $t-1$ days.

likely to be correlated over time within individuals, we calculate robust standard errors allowing for clustering at the individual level. Since our correction of the coding of workers' ages slightly changes the sample, we estimate a number of regressions using age cutoffs from 42 to 48.

The coefficient on the interaction of female and 28 days is not statistically significant at conventional levels for any of these specifications (Table 1, Panel A). The interaction coefficient equals 1.13 and is statistically significant at the 10 percent level when the age cutoff rises to 47 and 48 years old. However, the notion that the inclusion of older women should increase precision does not align with the sharp increase in variation of menstrual cycles immediately prior to menopause. This increased variability is shown in Figure 4, based on the data collected by Treloar et al. (1967) on the menstrual experiences of more than 2,700 women over a 30 year period. While cycle length and variability decline between the ages of 20 and 40, variability increases substantially as women move above the age of 40.

In addition to an indicator for female and an interaction between female and 28 days, IM include an interaction between female and any multiple of 7 days. The motivation for its inclusion would be that if women are also more likely to be absent on multiples of 7 for other reasons (e.g., weekly child care commitments), this control variable will mitigate bias in favor of the menstrual cycle hypothesis. However, we see no strong a priori reason to include this term, and find little evidence to support its inclusion. Specifically, we estimate specifications that include interactions of an indicator for female with indicators for every multiple of 7 up to 70 (Appendix Table 1), and the hypothesis of a single 7-day interaction term is easily rejected.[6] In Table 1 Panel B, we present the results of specifications that do not include an interaction of

---

[5] Additionally, as with our reproduction of Figure 1, we combine spells that form a continuous period of absence, so the distance between spells can never equal one day.

[6] We stop at 70 because the interaction term for multiples of seven used by IM only included values up to 70, and we see no need to go further.

female with multiples of 7 days. In all cases, the coefficients are slightly smaller than before and far from statistically significant.

We conclude from this analysis that the identification of an effect of menstruation on female absences based on 28 day cycles is quite weak. However, while this strategy is appealing, it is unclear whether one would expect to uncover anything given that less than 16% of menstrual intervals are exactly 28 days (Hartman (1972), Chiazze et al. (1968)). Indeed, in one of the most extensive studies on cycle length (Treloar et al. (1967)) the authors conclude:

> *"There is no substantial justification for the widely held belief that women normally vary in menstrual interval about a value of 28 days common to all. Each woman has her own central trend and variation, both of which change with age."*

The medical evidence on menstrual cycles suggests that one might look for significant increased absence in cycles around 28 days rather than only at 28. However, we present evidence that interpreting such estimates as informative about the menstrual cycle is problematic.

In Panel A of Table 2, we present estimates from hazard models that remove the interaction of female with 28 days and, instead, include interactions between female and ranges of days around 28 (i.e., 27-29, 26-30, or 25-31). For age cutoffs of 45 and above, these estimates are positive, smaller than the estimates for 28 days alone, and not statistically significant. For younger age cutoffs, the point estimates rise slightly, and, for the 26-30 day range, approach statistical significance at age cutoffs 42 and 43.

Nevertheless, we would argue that this is weak evidence in favor of the hypothesis that female absences are following the menstrual cycle. In support of this argument, we present the results of hazard regressions that use the same specification to contrast absence patterns of younger men with older men (Table 2, Panel B). Here, we find evidence that older men also show an increased hazard of absence in the range around 28 days, and the coefficients are actually larger and more precisely estimated than those contrasting younger females with younger males.[7] It is hard to believe that the menstrual cycle is influencing differential absence patterns by age among men, and this, we believe, makes it difficult to conclude that menstrual cycles play an important role in the differential absence patterns by gender among younger employees.

3. Discussion and Conclusion

We believe that there is no compelling evidence that female absences due to the menstrual cycle explain significant portions of the gender gaps in absences or earnings.[8] However, the lack of evidence for a biological explanation of the gender earnings gap does not imply that that menstruation never causes women to be absent from work. For example, premenstrual dysphoric disorder—estimated to affect roughly five percent of women ages 18 to 48 (Katz el al. (2007))—is associated with severe psychological/emotional symptoms that could easily lead women to miss work. Moreover, while women in Italy (or the U.S.) do not typically receive explicit support if they wish to remain home from work during menstruation, 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. Given the potential importance of settings and institutions, more research is needed to determine the extent to which biological gender differences, such as the menstrual cycle, may explain gender gaps in absences and earnings in different countries.

[7] To see more generally what drives this result, we plot the frequency distributions of distances between consecutive absences for men under 45 and men aged 45 and over in Figure 5. One can see a clear increase in the difference between the two distributions around 28 days.

[8] In a separate, earlier paper, posted online as a comment in the American Economic Journals discussion forum, we analyze much larger data set covering absences of full-time teachers in New York City public schools. We select our sample of teachers to match the restrictions imposed by IM for their sample of Italian bank employees (e.g., we exclude women who have gone on maternity leave). We find no evidence of an increase is the hazard rate of absence for female teachers at or around 28 days.

However, we are doubtful that the role of the menstrual cycle is likely to be accurately assessed with data on absences alone. The substantial variation in the length of menstrual cycles, both within and between women, greatly weakens the possibility of uncovering patterns of absence, such as a spike at 28 days, which could be clearly linked to the menstrual cycle. In addition to variation in cycle length, there is likely variation in the timing of symptoms, such as those due to premenstrual syndrome (PMS). While we have been unable to locate data on the variation in timing of PMS symptoms, they can potentially occur during any point in the second half of the menstrual cycle (Katz et al. (2007)).

A different approach would be to use data on both menstrual cycles and absences. Oster and Thornton (2009) use such data to document the role of menstruation in determining school attendance among girls in developing countries. We believe that this presents a more promising path for future research on whether menstruation plays an important role in determining gender differences in labor market outcomes.

References

Altonji, J.G. and Blank, R.M. (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, F.D. and Kahn, L.M. (2000) "Gender Differences in Pay," *Journal of Economic Perspectives* 14(4): 75-99.

Chiazze, L., Brayer, F.T., Macisco, J.J., Parker, M.P., Duffy, B.J. (1968) "The Length and Variability of the Human Menstrual Cycle," *Journal of the American Medical Association*. 203(6): 377-380.

Goldin, C. (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.

Hartman, C.G. (1972) *Science and the Safe Period; a Compendium of Human Reproduction*. Huntington, N.Y.: Robert E. Krieger Publishing Company.

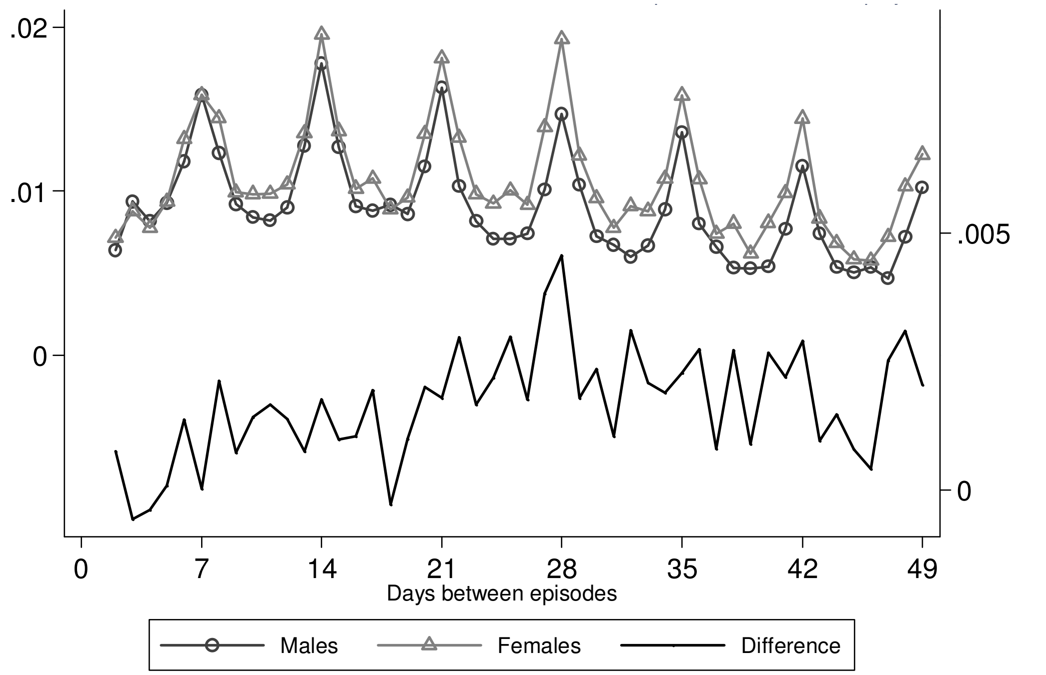
Ichino, A. and Moretti, E. (2009) "Biological Gender Differences, Absenteeism and the Earnings Gap," *American Economic Journal: Applied Economics* 1(1): 183-218

Katz, V.L., Lentz, G.M., Lobo, R.A., and Gershenson, D.M. (2007) *Comprehensive Gynecology, 5th Edition.* Philadelphia: Mosby, Elsevier.

Oster, E. and Thornton, R. (2009) "Menstruation and Education in Nepal," NBER Working Paper 14853

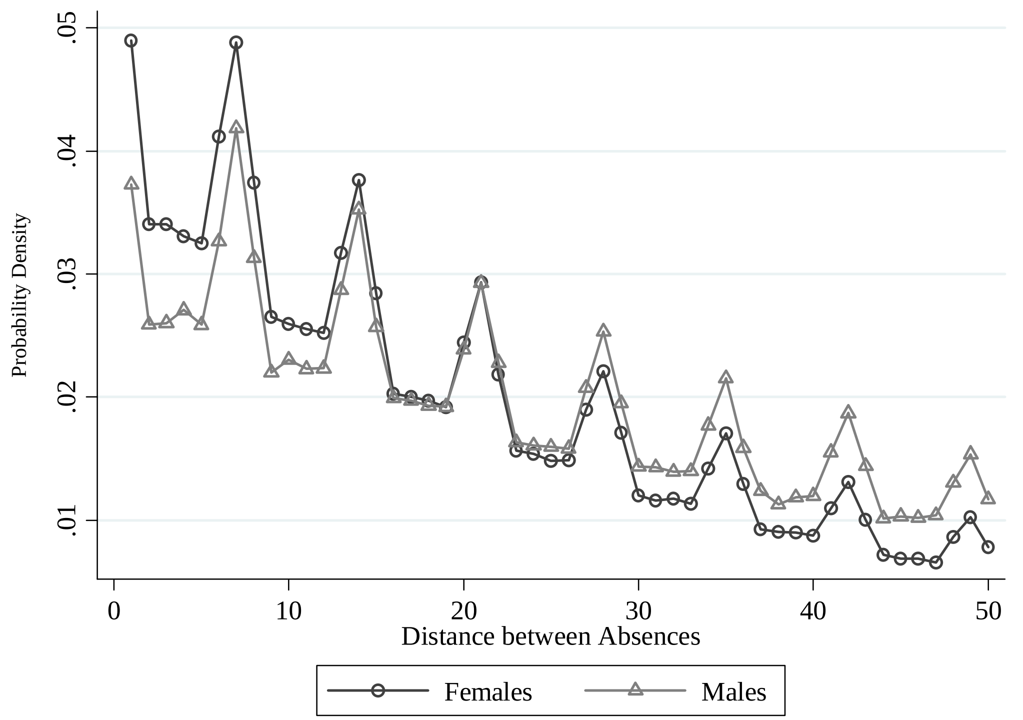
Treloar, A.E., Boynton, R.E., Behn, B.G., Brown, B.W. (1967) "Variation of the Human Menstrual Cycle through Reproductive Life." *International Journal of Fertility*. 12(1 Pt 2):77-126.

Figure 1: Distribution of the Distance between Consecutive Absence Spells



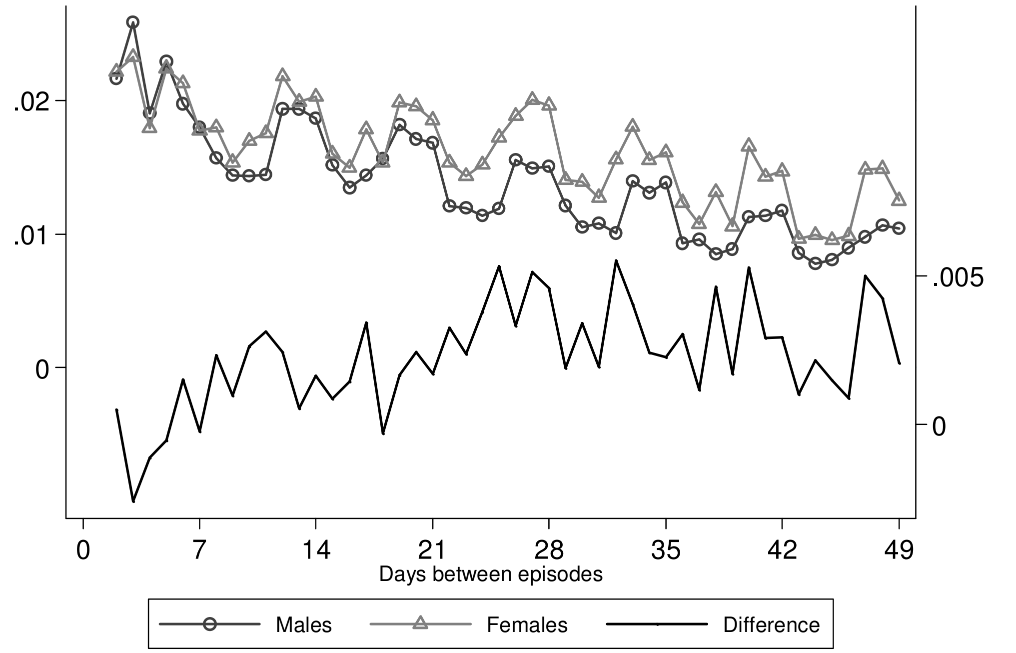
Notes: We display the frequency distribution of the distance in days between consecutive absence episodes for male and female Italian Bank employees (using the y-axis on the left side), as well as the difference between the two distributions (using the y-axis on the right side). The difference between the distributions replicates Figure 1 in Ichino and Moretti (2009), with the alteration that we do not rescale the male-female difference to have a mean of 0 and we combine spells that form a continuous period of absence, so the distance between spells can never equal one day.

Figure 2: Distance Between Consecutive Absences for Simulated Data



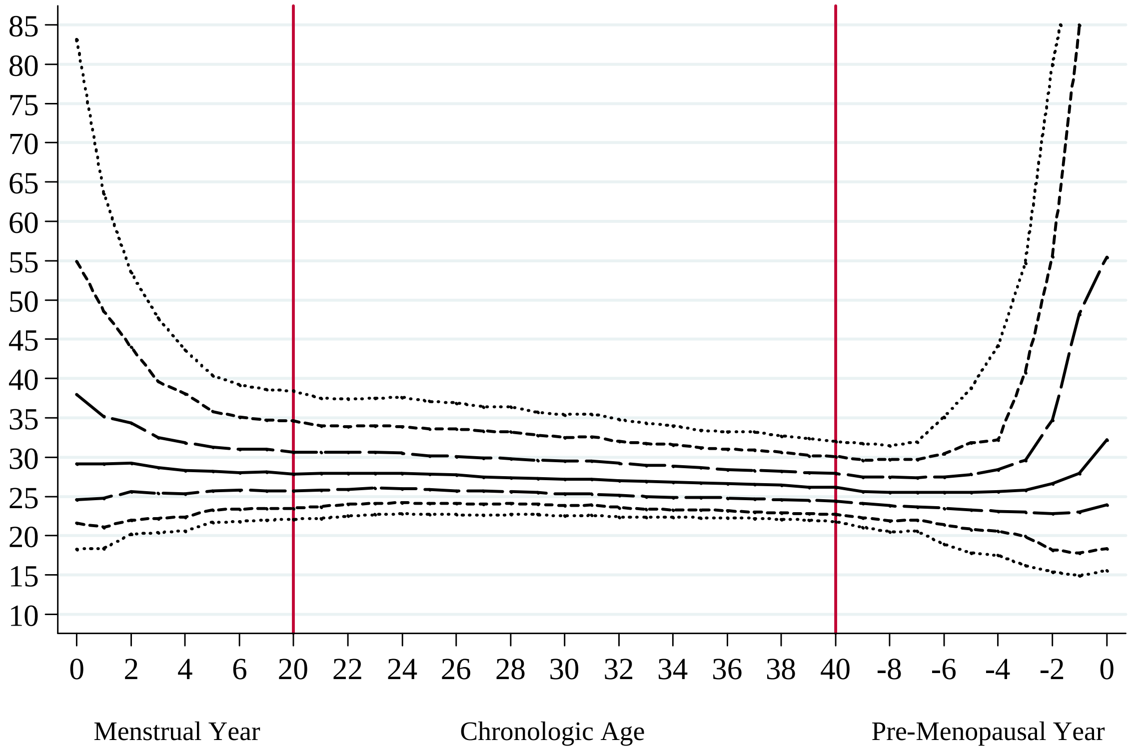
Note: This figure plots the distribution of distance in days between absences from simulated data. The simulation generates 25,000 employees of each gender and follows them for 1,000 calendar days. Females are absent on work days (Monday through Friday) with a 5 percent probability; males are absent on work days with a 3 percent probability.

Figure 3: Distribution of the Distance between Consecutive Absence Spells
Adjusted for Weekends and Length of Prior Spell



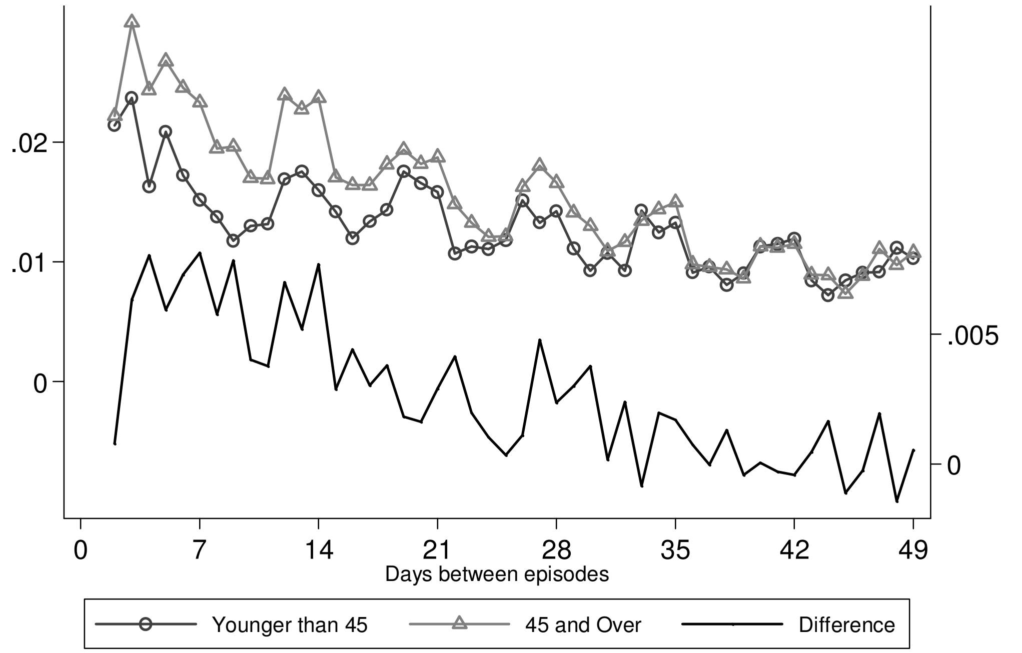
Notes: This figure display the frequency distribution of the distance in days between consecutive absence episodes for male and female Italian Bank employees (using the y-axis on the left side), as well as the difference between the two distributions (using the y-axis on the right side). Weekends and periods when a worker was still absent on a prior spell do not count towards the calculation of the frequency distribution.

Figure 4: Contours of Menstrual Cycle Length Frequency Distribution



Source: Treloar et al. (1967), Table 3. Menstrual year indicates time since menarche, the first occurrence of menstruation; pre-menopausal year indicates time until the woman reaches menopause. The lines from bottom to top are the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles. The mean age of menarche is 13, and the mean age of menopause is 51 (Katz et al. 2007).

Figure 5: Distribution of the Distance between Consecutive Absence Spells
Males Only, Adjusted for Weekends and Length of Prior Spell



Notes: This figure displays frequency distributions of the distance in days between consecutive absence episodes for older and younger male Italian Bank employees (using the y-axis on the left side), and the difference between the two distributions (using the y-axis on the right side). Weekends and periods when a worker was still absent on a prior spell do not count towards the calculation of the frequency distribution.

Table 1: Hazard of an Absence for Females Relative to Males and Risk of "Menstrual Cycle"

	IM (2009)	Re-analysis of IM data						
Panel A: Basic Results	Under 45	Under 42	Under 43	Under 44	Under 45	Under 46	Under 47	Under 48
Female	1.39	1.49	1.49	1.48	1.47	1.47	1.46	1.45
	(35.94)	(19.1)	(19.6)	(20.1)	(20.4)	(20.7)	(20.7)	(20.5)
Female*28 Days	1.15	1.08	1.08	1.10	1.11	1.07	1.13	1.13
	(2.16)	(1.0)	(1.0)	(1.3)	(1.4)	(1.0)	(1.7)	(1.8)
Female*Multiple of 7 Days	0.95	0.98	0.98	0.98	0.98	0.97	0.95	0.95
	(-2.04)	(-0.6)	(-0.6)	(-0.6)	(-0.6)	(-1.2)	(-1.9)	(-1.7)
# Observations (Days at Risk)	*n/a*	6,897,879	7,503,078	8,119,142	8,728,534	9,293,047	9,805,947	10,200,000
Panel B: Adjusted Specification		Under 42	Under 43	Under 44	Under 45	Under 46	Under 47	Under 48
Female		1.48	1.49	1.48	1.47	1.47	1.45	1.45
		(18.4)	(19.0)	(19.4)	(19.6)	(19.8)	(19.7)	(19.6)
Female*28 Days		1.06	1.06	1.08	1.09	1.04	1.08	1.09
		(0.8)	(0.8)	(1.2)	(1.2)	(0.6)	(1.1)	(1.2)
# Observations (Days at Risk)		6,897,879	7,503,078	8,119,142	8,728,534	9,293,047	9,805,947	10,200,000

Note: The first column of Panel A displays results from Ichino and Moretti (2009), Table 2. The remaining columns of Panel A display results from hazard regressions that use the same Italian Bank data but differ in that: (a) issues of right censoring, day of the week coding, and age coding have been corrected and (b) standard errors allow for clustering at the individual worker level. Panel B shows results of specifications that remove the interaction of female with multiples of 7 days (up to 70). Additional controls include an indicator for female, worker characteristics (see text for details) and day of week indicators. T-statistics are shown in parentheses.

Table 2: Hazard of an Absence in Ranges Around 28 Days

Panel A: Absence and Gender	Under 42	Under 43	Under 44	Under 45	Under 46	Under 47	Under 48
Female*27-29 Days	1.10	1.08	1.07	1.06	1.04	1.05	1.05
	(1.9)	(1.5)	(1.5)	(1.4)	(0.8)	(1.1)	(1.1)
Female*26-30 Days	1.10	1.09	1.07	1.07	1.05	1.06	1.05
	(2.3)	(2.0)	(1.8)	(1.6)	(1.2)	(1.4)	(1.3)
Female*25-31 Days	1.07	1.06	1.06	1.06	1.05	1.06	1.05
	(1.8)	(1.7)	(1.6)	(1.6)	(1.4)	(1.6)	(1.4)
# Observations (Days at Risk)	6,303,546	6,897,879	7,503,078	8,119,142	8,728,534	9,293,047	9,805,947

Panel B: Placebo (Males Only)	Age Split at 42	Age Split at 43	Age Split at 44	Age Split at 45	Age Split at 46	Age Split at 47	Age Split at 48
Older Male*27-29 Days	1.25	1.22	1.23	1.21	1.15	1.20	1.19
	(5.1)	(4.5)	(4.6)	(4.2)	(3.1)	(3.9)	(3.5)
Older Male*26-30 Days	1.24	1.23	1.23	1.20	1.16	1.20	1.17
	(5.7)	(5.3)	(5.3)	(4.7)	(3.7)	(4.4)	(3.6)
Older Male*25-31 Days	1.16	1.14	1.15	1.14	1.12	1.16	1.13
	(4.3)	(3.8)	(4.0)	(3.7)	(3.2)	(3.9)	(3.0)
# Observations (Days at Risk)	9,967,691	9,967,691	9,967,691	9,967,691	9,967,691	9,967,691	9,967,691

Note: Regressions in Panel A are limited to male and female workers under a certain age cutoff (defined for each column) and each cell displays a coefficient from an individual hazard regression on an interaction of being female with a range of days since prior absence. Regressions in Panel B include only males, and each cell displays the coefficient from an individual hazard regression on an interaction of being over a certain age (splits defined for each column) with a range of days since prior absence. Additional controls include an indicator for female, worker characteristics (see text for details) and day of week indicators. T-statistics (shown in parentheses) allow for clustering at the individual level.

Appendix Table 1: Hazard of an Absence for Females on Multiples of 7 Days

	Under 42	Under 43	Under 44	Under 45	Under 46	Under 47	Under 48
Female	1.49	1.49	1.48	1.47	1.47	1.46	1.45
	(19.1)	(19.7)	(20.1)	(20.4)	(20.7)	(20.7)	(20.5)
Female*7 Days	0.84	0.82	0.81	0.78	0.79	0.76	0.76
	(-1.7)	(-2.0)	(-2.2)	(-2.7)	(-2.6)	(-3.0)	(-3.0)
Female*14 Days	0.98	1.01	0.99	0.97	0.94	0.89	0.89
	(-0.3)	(0.1)	(-0.2)	(-0.4)	(-0.9)	(-1.7)	(-1.8)
Female*21 Days	0.81	0.84	0.83	0.83	0.84	0.84	0.83
	(-2.7)	(-2.4)	(-2.7)	(-2.7)	(-2.6)	(-2.7)	(-3.0)
Female*28 Days	1.06	1.06	1.08	1.09	1.04	1.07	1.08
	(0.8)	(0.8)	(1.1)	(1.2)	(0.5)	(1.0)	(1.1)
Female*35 Days	1.00	1.04	1.01	1.00	0.96	0.93	0.94
	(0.0)	(0.5)	(0.2)	(0.0)	(-0.6)	(-1.1)	(-0.9)
Female*42 Days	1.02	1.01	1.03	1.05	1.04	1.03	1.04
	(0.3)	(0.1)	(0.4)	(0.7)	(0.6)	(0.4)	(0.6)
Female*49 Days	1.03	1.04	1.05	1.10	1.11	1.13	1.12
	(0.4)	(0.5)	(0.6)	(1.2)	(1.4)	(1.6)	(1.5)
Female*56 Days	0.93	0.96	1.01	1.08	1.08	1.07	1.12
	(-0.8)	(-0.4)	(0.1)	(0.9)	(0.9)	(0.9)	(1.4)
Female*63 Days	1.19	1.20	1.21	1.18	1.16	1.10	1.10
	(1.8)	(2.0)	(2.1)	(1.9)	(1.7)	(1.1)	(1.2)
Female*70 Days	1.23	1.25	1.24	1.23	1.22	1.21	1.19
	(2.0)	(2.3)	(2.3)	(2.2)	(2.2)	(2.2)	(2.0)
Test on Equality of Coefficients for 7 day Multiples Excluding 28 (P-Value)	0.0386	0.0102	0.0074	0.0046	0.0035	0.0013	0.0011
# Observations (Days at Risk)	6,897,879	7,503,078	8,119,142	8,728,534	9,293,047	9,805,947	10,200,000

Note: Additional controls include an indicator for female, worker characteristics (see text for details) and day of week indicators. T-statistics (shown in parentheses) are calculated allowing for clustering at the individual worker level.