Appendix: Impact of 5-Day Workweek on Hazard at 7-Day Intervals

In the main text of our paper, we explain how the 5-day work week creates 7-day periodicity in absence spells that confound the estimation of differences in hazard rates at multiples of 7, including 28 days. We also show that including an interaction between a group indicator and an indicator for a multiple of 7 in a hazard regression will not necessarily solve the problem if groups differ in the persistence of absence spells or recovery rates.

A potential alternative solution we explore in this appendix is to analyze data that have been converted to a 5-day (work) week. If our prior is that group hazards should differ at 28 calendar days, then we might test whether there is a spike the difference in group hazard rates at 20 working day intervals. We show here that, like the use of a constant interaction term, the validity of this test depends on the data generating process (DGP) for absences. If the data generating process is extremely simple then this method should be able to distinguish an excess hazard for a particular group at 28 days, but this method is still confounded by the 7-day periodicity under a slightly more complex (and more realistic) DGP.

Consider a very simple DGP where absences occur randomly on weekdays with probability 7 percent for females and 5 percent for males, as discussed in Section 2 of our paper. Kaplan-Meier hazard rates for simulated data on absences for a sample of 10,000 females and 10,000 males over 1,000 days based on this DGP are shown in Figure A1 (left panel), as well as the gender difference in hazard rates (right panel). These replicate Figure 1 from our paper and, as we explain there, show 7-day periodicity despite no assumption of differential probabilities of absence by day of the week (e.g., no “Monday Effect”). However, rescaling the data to a 5-day workweek reduces periodicity in absences (Figure A2).

Figure A1: Simple DGP Hazards, 7-Day Week

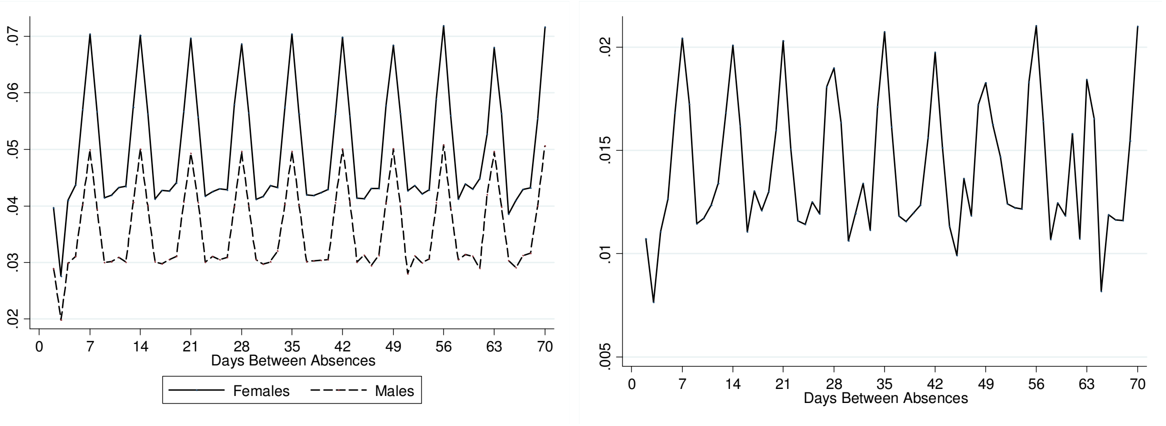
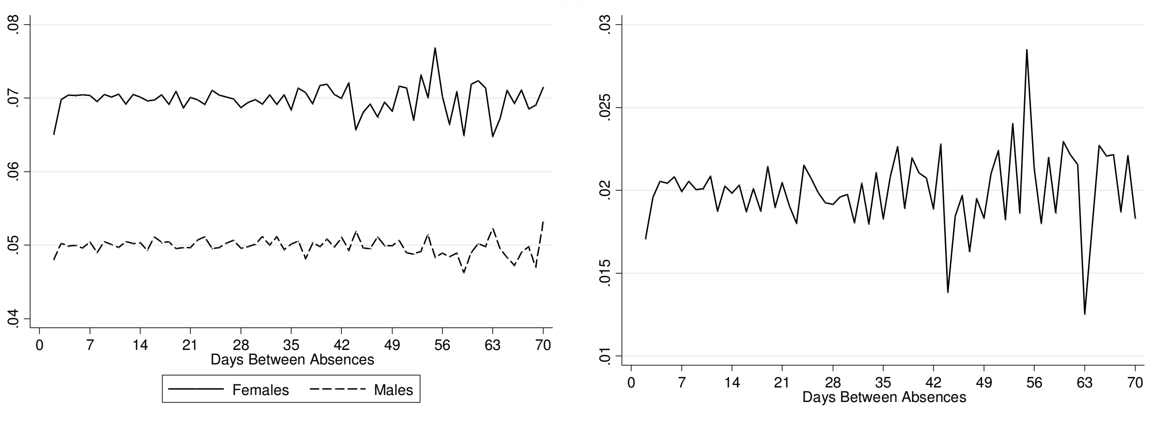


Figure A2: Simple DGP Hazards, 5-Day Week



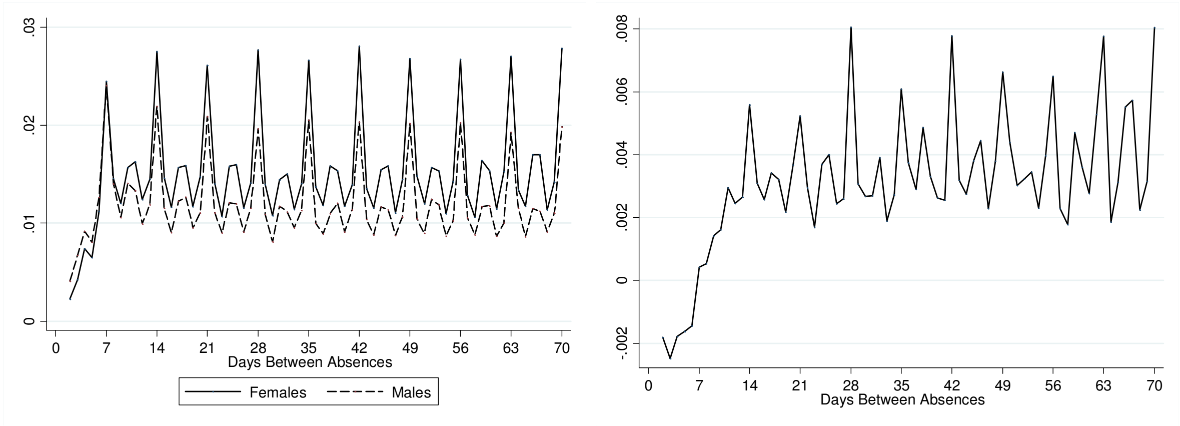
However, the DGP does not allow health shocks to be persistent or for individuals to delay returning to work until they are well enough so that an immediate remission to absence is unlikely. We incorporate these modestly more complicated features into a DGP by assuming that health for individual *i* on day *t* follows an AR-1 process:

$$H_{i,t} = \rho_g H_{i,t-1} + \varepsilon_{i,t}$$

where ρ_g measures the (possibly gender-specific) persistence of health shocks. We assume there are two (gender-specific) thresholds $c_{g,a}$ and $c_{g,p}$, where $c_{g,a} < c_{g,p}$. Individuals become absent if $H_{i,t} < c_{g,a}$ and, once absent, return to work when $H_{i,t} > c_{g,p}$. This DGP replicates the empirical facts that individuals are more likely to be absent on day *t* if they are absent on day *t-1*, and that individuals' spells of absence do not usually occur in quick succession. The earlier model can actually be thought of a special case of this, where $\rho_g = 0$ and $c_{g,a}= c_{g,p}$. As before, individuals cannot be absent on weekends.

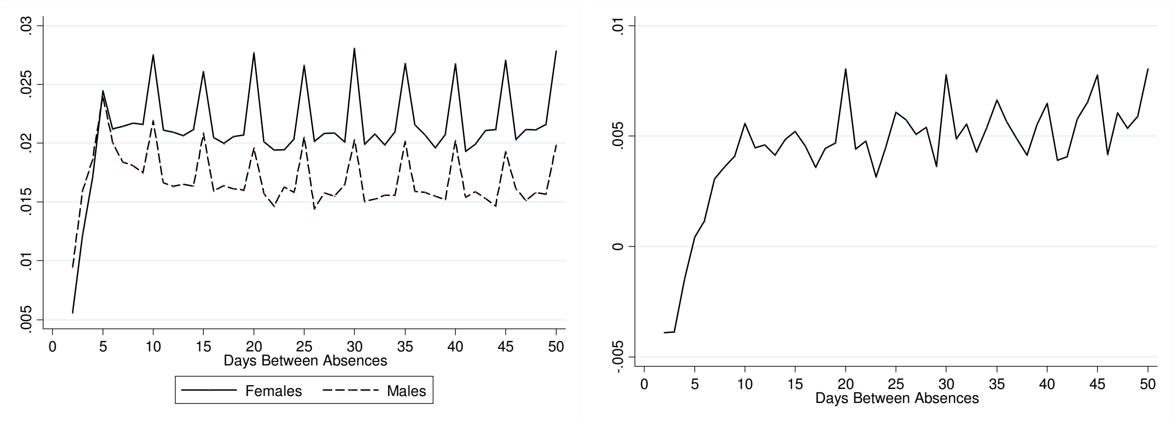
Again, we simulate a sample of absences for 10,000 females and 10,000 males over 1,000 days. We choose parameters that generate workday absence rates of about 5 percent for men and 7 percent for women: $H_{i,0} \sim N(0,1)$, $\varepsilon_{i,t} \sim N(0,1)$, $\rho_f = 0.75$; $\rho_m = 0.78$; $c_{m,a} = -3.1$; $c_{f,a} = -2.8$; $c_{m,p} = -1.6$; $c_{m,a} = -1$. As we show in Figure 1 of our paper, there are positive spikes in the hazards for both genders at multiples of 7, and, more importantly, there are large, non-constant, differences between the female and male hazards at different multiples of 7. We replicate this here as Figure A3. Coincidentally, the spike at 28-days is the greatest in this example, despite the fact that we have made no assumption that women should be more likely to have spells 28 days apart.

Figure A3: DGP with Persistence and Recovery, 7-Day Week



Importantly, rescaling to a 5-day workweek does not fix this problem (Figure A4). We still find spikes and group differences in hazard rates at multiples of 5. In this simulation, the largest spike in group difference happens to come at 20 (work) days, and could thus provide spurious evidence in favor of 28-day cycles in absences for females.

Figure A4: DGP with Persistence and Recovery, 5-Day Week



Appendix Table 1: Hazard of an Absence for Females Relative to Males - Italian Bank Data, Various Age Cut-offs

Panel A	Under 42	Under 43	Under 44	Under 45	Under 46	Under 47	Under 48	Under 49	Under 50	Under 51	Under 52	Under 53	Under 54
Female	1.49	1.49	1.48	1.47	1.47	1.46	1.45	1.45	1.45	1.45	1.45	1.44	1.44
	(19.14)	(19.65)	(20.12)	(20.38)	(20.67)	(20.69)	(20.54)	(20.50)	(20.65)	(20.78)	(20.72)	(20.71)	(20.68)
Female*28 Days	1.08	1.08	1.10	1.11	1.07	1.13	1.13	1.13	1.14	1.14	1.13	1.12	1.12
	(0.99)	(0.96)	(1.30)	(1.39)	(0.96)	(1.69)	(1.76)	(1.71)	(1.96)	(1.91)	(1.80)	(1.71)	(1.73)
Female*Multiple of 7 Days	0.98	0.98	0.98	0.98	0.97	0.95	0.95	0.96	0.96	0.96	0.96	0.96	0.95
	(-0.61)	(-0.60)	(-0.63)	(-0.63)	(-1.19)	(-1.86)	(-1.72)	(-1.48)	(-1.60)	(-1.49)	(-1.59)	(-1.69)	(-1.83)
Panel B	Under 42	Under 43	Under 44	Under 45	Under 46	Under 47	Under 48	Under 49	Under 50	Under 51	Under 52	Under 53	Under 54
Female	1.49	1.49	1.48	1.47	1.47	1.46	1.45	1.45	1.45	1.45	1.45	1.44	1.44
	(19.15)	(19.65)	(20.12)	(20.38)	(20.67)	(20.70)	(20.55)	(20.50)	(20.66)	(20.79)	(20.72)	(20.71)	(20.69)
Female*28 Days	1.06	1.06	1.08	1.09	1.04	1.07	1.08	1.08	1.10	1.09	1.08	1.07	1.07
	(0.80)	(0.77)	(1.12)	(1.21)	(0.54)	(1.02)	(1.15)	(1.20)	(1.40)	(1.40)	(1.24)	(1.10)	(1.07)
Female*7 Days	0.84	0.82	0.81	0.78	0.79	0.76	0.76	0.75	0.74	0.75	0.76	0.76	0.75
	(-1.72)	(-2.03)	(-2.25)	(-2.71)	(-2.64)	(-3.01)	(-3.02)	(-3.20)	(-3.42)	(-3.34)	(-3.30)	(-3.28)	(-3.39)
Female*14 Days	1.01	0.99	0.97	0.94	0.89	0.89	0.89	0.89	0.89	0.89	0.88	0.87	0.86
	(0.11)	(-0.16)	(-0.39)	(-0.88)	(-1.73)	(-1.80)	(-1.81)	(-1.75)	(-1.77)	(-1.82)	(-2.01)	(-2.16)	(-2.45)
Female*21 Days	0.84	0.83	0.83	0.84	0.84	0.83	0.83	0.85	0.85	0.86	0.86	0.86	0.86
	(-2.39)	(-2.67)	(-2.72)	(-2.60)	(-2.73)	(-3.03)	(-3.05)	(-2.58)	(-2.64)	(-2.52)	(-2.60)	(-2.56)	(-2.64)
Female*35 Days	1.00	1.04	1.01	1.00	0.96	0.93	0.94	0.95	0.95	0.94	0.95	0.96	0.96
	(0.00)	(0.51)	(0.15)	(0.00)	(-0.56)	(-1.10)	(-0.94)	(-0.83)	(-0.81)	(-1.02)	(-0.90)	(-0.70)	(-0.69)
Female*42 Days	1.02	1.01	1.03	1.05	1.04	1.03	1.04	1.04	1.05	1.06	1.06	1.05	1.07
	(0.30)	(0.11)	(0.43)	(0.70)	(0.58)	(0.43)	(0.56)	(0.58)	(0.73)	(0.92)	(0.84)	(0.83)	(1.06)
Female*49 Days	1.03	1.04	1.05	1.10	1.11	1.13	1.12	1.13	1.11	1.12	1.11	1.09	1.09
	(0.36)	(0.51)	(0.60)	(1.17)	(1.39)	(1.58)	(1.51)	(1.65)	(1.46)	(1.55)	(1.43)	(1.28)	(1.19)
Female*56 Days	0.93	0.96	1.01	1.08	1.08	1.07	1.12	1.12	1.15	1.16	1.15	1.16	1.14
	(-0.81)	(-0.43)	(0.09)	(0.89)	(0.94)	(0.87)	(1.40)	(1.50)	(1.85)	(1.91)	(1.92)	(1.97)	(1.78)
Female*63 Days	1.19	1.20	1.21	1.18	1.16	1.10	1.10	1.09	1.08	1.09	1.09	1.08	1.10
	(1.77)	(1.99)	(2.10)	(1.92)	(1.75)	(1.14)	(1.21)	(1.07)	(0.94)	(1.14)	(1.15)	(1.05)	(1.30)
Female*70 Days	1.23	1.25	1.24	1.23	1.22	1.21	1.19	1.22	1.21	1.20	1.21	1.18	1.18
	(2.04)	(2.31)	(2.31)	(2.23)	(2.22)	(2.17)	(1.96)	(2.35)	(2.26)	(2.18)	(2.26)	(2.03)	(2.01)

Note: Each column displays results from hazard regressions that use the Italian Bank data where issues of one-day adjacent spells, right censoring, day of the week coding, and age coding have been corrected and standard errors allow for clustering at the the individual worker level. In Panel B, the interaction of female with multiples of seven has been omitted from the specification. All specifications control for age, years of schooling, marital status, number of children, managerial occupation, seniority, and dummies for each day of the week. T-statistics are shown in parentheses. A hazard ratio of 1 indicates no effect.