DISCUSSION PAPER SERIES

IZA

IZA DP No. 7480

Gender Differences in Preferences for Health-Related Absences from Work

Daniel Avdic
Per Johansson

June 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Gender Differences in Preferences for Health-Related Absences from Work

Daniel Avdic
Uppsala University,
UCLS and IFAU

Per Johansson
Uppsala University,
UCLS, IFAU and IZA

Discussion Paper No. 7480
June 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 7480
June 2013

ABSTRACT

Gender Differences in Preferences for Health-Related Absences from Work*

Women are on average more absent from work for health reasons than men. At the same time, they live longer. This conflicting pattern suggests that part of the gender difference in health-related absenteeism arises from differences between the genders unrelated to actual health. An overlooked explanation could be that men and women's preferences for absenteeism differ, for example because of gender differences in risk preferences. These differences may originate from the utility-maximizing of households in which women's traditional dual roles influence household decisions to invest primarily in women's health. Using detailed administrative data on sick leave, hospital visits and objective health measures we first investigate the existence of gender-specific preferences for absenteeism and subsequently test for the household investment hypothesis. We find evidence for the existence of gender differences in preferences for absence from work, and that a non-trivial part of these preference differences can be attributed to household investments in women's health.

JEL Classification: I13, J22, D13

Keywords: sickness absence, gender norms, health investments

Corresponding author:

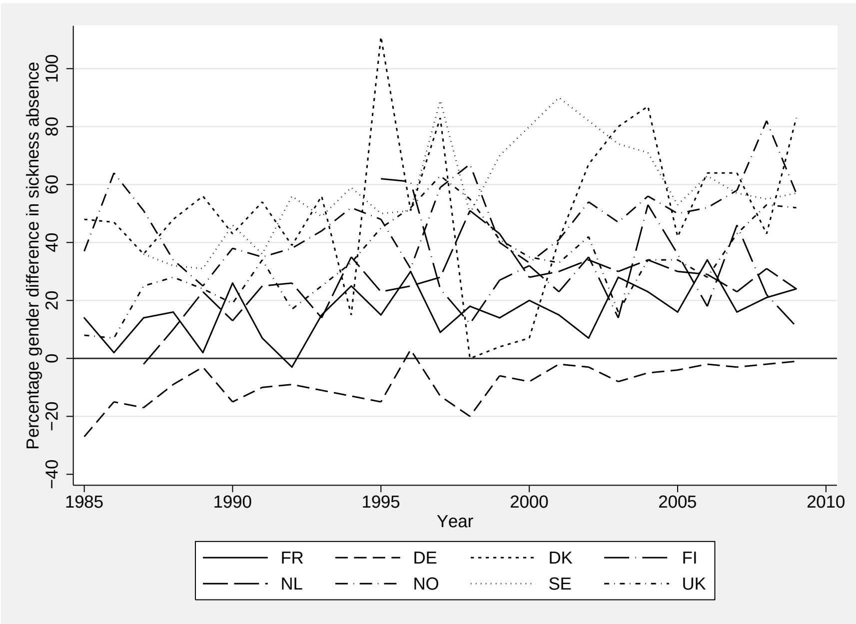
Per Johansson
The Department of Economics
Uppsala University
751 20 Uppsala
Sweden.
E-mail: per.johansson@ifau.uu.se

* We are grateful for helpful comments from Nicolas Ziebarth, Johan Vikström, Will White and seminar participants at the 2012 ESPE conference in Bern, the 2012 EALE conference in Bonn, the Department of Economics and the Health Economic Forum at Uppsala University and the 15th IZA Summer School in Labor Economics in Munich. Financial support from the Swedish Council for Working Life and Social Research (DNR 2004-2005 and 2009-0826) is gratefully acknowledged.

1 Introduction

Health-related absenteeism has been persistently higher for women than for men in most developed countries over the last decades. For example, *Figure 1.1* illustrate the percentage gender difference in reported sickness absence for a number of European countries over time. Similar gender patterns of absence emerge regardless of whether one look at the extensive or intensive margin of absenteeism or whether the data originates from surveys or administrative registers (Mastekaasa and Olsen, 1998). Morover, these observations are also in line with observed gender differences in many morbidity measures such as medical care utilization and measures of self-reported health (Sindelar, 1982).

Figure 1.1: Gender difference in the prevalence of sickness absence by country, 1980-2010



NOTE.—The data is provided by Eurostat. The vertical axis is defined as the percentage difference in the share of women divided by the share of men that reported absence from work for health reasons at some time during a specific period of time.

If one were to use health-related absenteeism (sick leave) as an objective measure of health, one would come to the conclusion that women have poorer health than men. At the same time, women live longer than men. In fact, the remaining life expectancy is longer for women than for men at all ages and in nearly all parts of the world. The global average gender difference in life expectancy was about four years in 2010 (Lee, 2010). Moreover, this gender pattern in mortality has also been very persistent over time.

A common explanation for these apparently contradictory observations are gender differences in health-related behavior such as lifestyle; and in particular the fact that women in general act more proactively than men (Nathanson, 1975; Verbrugge, 1982; Sindelar, 1982; Schappert and Nelson, 1999; Stronegger *et al.*, 1997; Uitenbroek *et al.*, 1996).[1] This explanation is supported by experimental evidence showing that women in general are more risk averse than men (see e.g. surveys in Eckel and Grossman (2008); Croson and Gneezy (2009); Bertrand (2010)). Hence, if women devote more attention to potential future illnesses, e.g. by utilizing medical services more frequently or by being absent from work more often (in order to recuperate) than men, poor health may be detected at an earlier stage, remediated and, as a consequence thereof, increase the life expectancy of women relative to men.

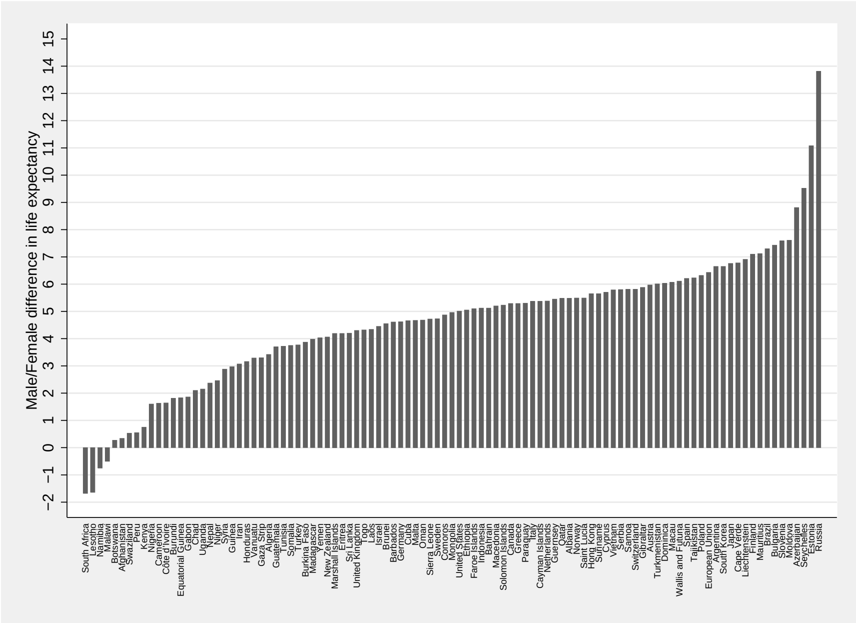
Why would men and women differ in their preferences towards risk and prevention? In principle, these differences may arise either from biological differences between the genders or from historical social processes forming perceptions and norms of what constitutes and distinguishes male and female behavior. Although the task of separating the effects of heredity and environment on social outcomes is a heroic one, the latter explanation is to some extent supported by the large cross-country variation in relative gender life expectancy around the world displayed in *Figure 1.2*. For example, the difference exceeds 14 years in Russia, is approximately five years in northern Europe, three in Asia and in South Africa men even outlive women by 1.5 years. Such large gender differences in life expectancy can hardly be attributed solely to biological differences between men and women in different geographical areas, but is also likely to contain elements of gender-specific health behaviors based on specific cultural norms in different parts of the

[1] This particular explanation for the so-called *morbidity-mortality* paradox was discussed already in the 17th century. The English demographer John Graunt elaborated on the observation that both the birth and death rates of men were higher than for women:

> "It appearing, that it were fourteen men to thirteen women [born], and that they die in the same proportion also, yet I have heard Physicians say, that they have two women Patients to one man, which Assertion seems very likely...Now, from this it should follow, that more women should die than men, if the number of Burials answered in proportion to that of Sicknesses; but this must be salved, either by the alledging, that the Physicians cure those Sicknesses, so as few more die, if none were sick; or else that men, being more intemperate than women, die as much by reason of their Vices, as women do by the Infirmitie of their Sex, and consequently, more Males being born then Females, more also die." (Graunt, 1662).

world.

Figure 1.2: Difference in Female-Male life expectancy by country, 2011



NOTE.—The data is provided by the CIA factbook, 2011. The vertical axis is measured as the difference of the female-male expected years of life for each country. The number of countries is a random sample of the total number of countries listed. The overall average among these countries is approximately four years.

The reasons for the gender gap in health-related absence have scarcely been investigated up until now, despite the geographical extent and vast policy implications on, for example, gender equality. The few studies that do exist in the field have focused on the roles that financial incentives, health and labor market conditions have played in explaining the gender differences.

Mastekaasa and Olsen (2000) examine whether the gender differences arise through segregation in the labor market in which, the authors claim, women mainly work in more unhealthy sectors. They find, however, that on the contrary gender differences in absenteeism actually increase when comparing men and women in the same occupation and labor market sector. Broström *et al.* (2004) and Angelov *et al.* (2011) use Swedish data to perform Oaxaca-Blinder decompositions of the gender absenteeism gap into differences in observed attributes and (unobserved) preferences. The main conclusion from these two studies is that economic incentives explain some of the differences while the working environment and health do not explain any of the differences. Interestingly, a substantial part of the differences could be explained by differences in the response to self-reported health.

Angelov *et al.* (2011) study the effects of parenthood on relative absenteeism and find that mothers increase their absence by 0.5 days per month more than fathers for as long as 18 years after the birth of the first child. Åkerlind *et al.* (1996) similarly investigate the relation between absenteeism, age and family status and find that young women and men with children utilize the sickness insurance scheme the most and the least respectively. Both studies conclude that gender inequality regarding the responsibility for child rearing seems to be a major factor behind the gender gap in absenteeism.

Laaksonen *et al.* (2008) use data from Finland and find that both short and medium absence spells (less than 60 days) are more common among women; a finding they contribute to a greater physical work load and work fatigue for women rather than psychosocial working conditions and family-related factors.

Finally Evans and Steptoe (2002) investigate the importance of socially constructed gender identities using survey data of men and women in occupations where they are in a minority (accounting for women and nurses for men). In particular, they investigate whether being the minority gender at the work place is associated with poorer health and mental well-being compared to being the majority gender. They find some empirical support for their hypothesis.

This paper contributes to the literature by first estimating differences in gender preferences for sickness absence and second, by investigating a particular reason why such a gender difference would arise. In particular, we estimate the relative effects on sickness absence for men and women of a negative health shock as measured by an observed hospital admission. Essentially, if the preferences for sickness absence differ between the genders, we expect the sickness absence response to hospitalization to be greater for women than for men, ceteris paribus.

We base our analysis on detailed Swedish longitudinal administrative data on hospital admissions to which we have merged administrative data on sickness absence, mortality and socioeconomic variables. The longitudinal data allows us to condition on pre-admission gender differences in health, economic incentives and other potential factors associated with health which may affect sickness absence in men and women differently.

We find that women increase their sickness absence more than men after a spell in hospital. This difference is remarkably stable across labor market sectors and different categories of diseases and is, in addition, also robust to a number of sensitivity analyses. Moreover, we find that the post-hospitalization mortality rates of women are lower than those of men. This allows us to rule out the possibility that the difference in post-admission sickness absence is driven by selective differences in the severances of the observed health shocks between the genders. Thus, all in all we find compelling evidence that men and women differ in their preferences for sickness absence or, put differently, health-related work absence.

Further, we examine whether the estimated gender differences in preferences for sickness absence arises from household responsibilities by studying whether the differences can mainly be explained by women with as opposed to without children (as a proxy for household responsibilities). The argument underlying this test is the empirical observation that a woman has the main responsibility for household production also when she works on the labor market, while her spouse mainly specializes in the labor market. The implication is that women with households have a dual role while their spouses only have one. At least two theories exist as to why women's dual roles may contribute to the observed gender differences in sickness absence: The first theory emphasizes the economic consequences of the division of labor in a household, while the other focuses on gender differences in health from the greater psychological pressure of multiple roles.

The first theory — denoted the *household health investment theory* — states that due to her dual roles, a woman's health is more important for the household than the health of her spouse. The reason is that her illness does not only include the lost earnings from being unable to perform market work, but it also creates an additional cost arising from her incapacity to perform home production (see Paringer (1983)). Therefore, it may in this context be rational for the household to respond more to a negative health shock of the woman (e.g. by starting a new or extending an already ongoing spell of sickness absence) than for an equivalent health shock of the spouse. In other words, it is more beneficial for the household to invest more in the woman's rather than in her spouse's health.[2]

[2]Note that we do not consider preferences as being stable or innate. This means that individuals without

In contrast to the theory of health investments, advocates of the second theory have emphasized that the gender gap in sickness absence may arise from the psychological pressure of dual roles — referred to here as the "double burden" of women (see e.g. Bratberg *et al.* (2002)).[3] Advocates of this *role strain theory* argue that multiple roles are detrimental for the well-being of the individual and may thus increase sickness absence. This could be the case, for example, if switching between the roles of performing tasks at work and at home entails a fixed cost, in particular for individuals where the roles are very different, for instance, working in an office and taking care of children.[4]

Our estimation results lend some support to the household health investment theory. In particular, the health investment effect can explain approximately one third of the previously estimated absence behavior difference between the genders.

The paper is organized as follows. The following section describes the relevant aspects of the Swedish sickness insurance system. Section three provides a stylized model framework from which we deduce empirically testable hypotheses regarding the absence behavior of men and women. Section four describes the data and the sampling method used for the estimation. Section five presents the results and section six offers some concluding remarks.

2 The Swedish sickness insurance system

All workers in Sweden, both the employed and unemployed, are covered by public health, sickness and disability insurance. The levels of compensation in the sickness and disability insurance are - in an international comparison - high (around 80 % and 65 %) and

families may have different preferences for, e.g., risk than individuals with families. Any observed difference between individuals with and without families could thus be from sorting (i.e. innate or biological) or the or simply formed in society or family.

[3]Time use studies in Sweden (SCB, 2009) have consistently shown that the total time worked is approximately the same for men and women. This similarity in total time worked corresponds well with statistics from time use studies in the USA, Germany and the Netherlands (Burda *et al.*, 2008). Hence, the double burden hypothesis should not be interpreted as an effect of a greater workload for women in comparison to men, but rather as an effect of the psychological strain of switching between roles.

[4]There is also literature advocating the potential benefits of having multiple roles (the *role enhancement* theory). According to this theory, individuals may feel that their lives are more meaningful when they perform several roles, which subsequently increase their well-being. Thus, the role enhancement theory gives exactly the opposite predictions of the role strain theory. See e.g. Mastekaasa (2000).

the degree of monitoring is lax (see e.g. Engström and Johansson (2012) for a recent detailed description of the institutions). Based on the information in a medical certificate, the Swedish Social Insurance Agency (SSIA) determines whether the illness in question leads to a reduced capacity for work (i.e. the inability to work). The proportion of cases where the SSIA decides against the physician's recommendation is, however, very small. For example, in 2006 the request for sick pay was rejected in 1.5 % of all new cases (SSIA, 2007:1). The length of a sick spell is to a large extent determined by the reasons given by the insured's own motivation (Arrelöv *et al.*, 2006). Moreover, physicians often make decisions against their better judgement by, for example, prescribing too long sickness absence spells (Englund, 2001). It is thus reasonable to assume that sickness absence is not only determined by objective health, but that there is also room for an individual's own judgment of his or her health.

3 Methodological framework

Our point of departure is the following stylized (Swedish) world: Market prices for household goods are high (except for highly subsidized child care) due to high minimum wages and high income taxes. The income tax system is based on individual rather than household income[5] and there is also a public sickness and disability insurance that replace earnings for individuals with poor health who are incapable of participating in the labor market. Individuals with children are assumed to live in households consisting of a man, a woman and at least one child. In contrast, individuals without children are assumed to be living alone.

Due to the individual-based income tax and costly household services, households would benefit in economic terms, under realistic assumptions of labor productivity (e.g. the existence of individual-specific learning-by-doing), if one of the household member are specializing in labor market production and the other to divide her/his time between both labor market and household production. We assume the former (specializing) mem-

[5]In 1971 Sweden changed from household taxation to individual income taxation. Selin (2009) estimated that female labor supply increased by 10 percentage points as a result of the reform.

ber to be the male and the latter (diversifying) to be the female.[6] We also assume that the time available for household production is increasing in health (see e.g. Grossman (1972) for a theoretical argument justifying this assumption). Finally, while productivity at the workplace is difficult to monitor, home production is not subject to this type of information asymmetry since the former usually involves working for someone else while the latter more resembles a form of self-employment. The implication is that shirking your duties is possible at the workplace but not in the household. These assumptions together suggest that the sickness absence incentives in a household are greater for women than for men for any given level of health.

Moreover, gender differences in absenteeism may arise from the highly segregated Swedish labor market (see e.g. SOU (2004:43)). For instance, if women predominantly work in sectors and/or establishments with poorer health-related working conditions than in more male-oriented sectors and establishments this might explain some of the gender difference.[7] These and several other factors complicate testing for behavioral differences between the genders.

Our approach to dealing with these difficulties is to estimate the average gender difference in the change of sickness absence from an adverse shock to individual health. Using the longitudinal features of the data we are able to condition on gender level differences in sickness absence due to confounding factors such as different economic incentives and labor market segregation. However, before we describe our empirical strategy in more detail we discuss a stylized theoretical model for the decision to be absent from work.

3.1 Theoretical framework

Let household preferences be represented by a direct utility function

$$u(C, \mathbf{s}, \mathbf{L}), \tag{1}$$

[6]The reason for this, admittedly, simplification of reality is that women, traditionally, perform most of the home production (see e.g. Boye (2008); Booth and Van Ours (2005); Evertsson and Nermo (2007); Tichenor (1999)).

[7]In general, however, we would expect the opposite. That is, men face poorer working conditions than women (see e.g. Broström *et al.* (2004)).

where C is the household's consumption of goods, $\mathbf{s} = (s_f, s_m)'$ represents the vector of desired daily hours of absence for the (f)emale and the (m)ale, and $\mathbf{L} = (L_f, L_m)'$ is the vector of contracted leisure hours for the spouses. Household services are produced by the female, h_f^p, and the male, h_m^p, but are jointly consumed. We assume that there are no productivity differences between the genders and normalize the price to one (that is, the value of consumed goods is normalized by the price for household services). Hence, the household production is $\mathbf{1}'\mathbf{h}^p$, where $\mathbf{h}^p = (h_f^p, h_m^p)$ and $\mathbf{1}$ =$(1,1)$.

The utility function is maximized under the following household budget constraint

$$\mathbf{w}'\mathbf{h} + y - (1-\delta)\mathbf{w}'\mathbf{s} - \mathbf{1}'\mathbf{h}^p = C, \tag{2}$$

where $\mathbf{w} = (w_f, w_m)$ is the vector of net wage rates, $\mathbf{h} = (h_m, h_f)$ are the contracted daily hours of work, $\delta \in (0,1)$ is the level of compensation in the sickness insurance and y is the family's non-labor income. Individuals also face the time constraint

$$T = L_j + h_j + h_j^p, \; j = f, m,$$

where $h_j = s_j + h_j^w$ with h_j^w as the desired daily hours of work.

Substitution of C from (2) into (1) and maximizing with respect to the absence of an arbitrary spouse, say the female, yields the household member's demand for absence conditional on both spouses' labor supply. Under the assumption of weak separability with respect to s_m, the solution can be written as

$$s_f = f(\mathbf{h}, \mathbf{h}^p, C_{sf}, \mu_f), \tag{3}$$

where $\mu_f = \mathbf{w}\mathbf{h} + y - (1-\delta)w_m s_m$ is the income net of sickness compensation (the virtual income) and $C_{sf} = \partial C / \partial s_f = -(1-\delta)w_f$ is the net cost of being absent from work. Due to the weak separability assumption the male absence time s_m only has an income effect through μ_f in (3). By simply changing s_m for s_f we obtain a corresponding expression for the male, i.e. for s_m, yielding $\mu_m = \mathbf{w}\mathbf{h} + y - (1-\delta)w_f s_f$ and $C_{sm} = -(1-\delta)w_m$. For singles we have that $\mu_j = w_j h_j + y$ and $h_j^p \equiv 0, \; j = f, m$.

An increase in the cost of absenteeism or virtual income would decrease and increase the demand for absence, respectively; hence $\partial s_j/\partial C_{sj}(=\phi_j) < 0$ and $\partial s_j/\partial \mu_j(=\zeta_j) > 0$. Moreover, it is reasonable to assume that increased hours in home production would increase demand for absence, hence $\partial s_j/\partial h_j^p(=\omega_j) > 0$. This last effect may arise for at least three reasons: First, (i) it may be an effect from reduced health because home production may impair health *per se*, i.e. the double burden hypothesis (see e.g. Bratberg *et al.* (2002)). Second, (ii) performing a greater degree of household production implies a greater potential payoff from responding to an illness at an early stage in order to reduce future lost home production, i.e. household investments in health (see e.g. Paringer (1983)). Finally, (iii) sickness absence does not necessarily prohibit some home production. Hence, in a context where monitoring is lax or difficult due to asymmetric information, sickness absence and home production can be regarded as substitutes.

Introducing heterogeneity in health, η_i (where a higher value of η_i implies better health), we get

$$s_{ij} = f(\eta_i, \mathbf{h}_i, \mathbf{h}_i^p, C_{isj}, \mu_{ij}).$$

It is reasonable to assume

$$\partial s_{ij}/\partial \eta_i(=\gamma_{ij}) < 0,$$

and

$$\partial^2 s_{ij}/\partial \eta_i \partial h_i^p(=\delta_{ij}) < 0.$$

In other words, better health reduces sickness absence in general but at an accelerating rate with respect to household production. Thus, if we could observe (exogenous) changes in health together with data on sickness absence and household production we could recover, or identify, the average demand for absenteeism of men and women, $\gamma_f = E(\gamma_{if}) < 0$ and $\gamma_m = E(\gamma_{im}) < 0$ with $\kappa = \gamma_f - \gamma_m < 0$, respectively, and for men and women with different levels of household production, $\delta_j = E(\delta_{ij}) < 0, j = m, f$.

3.2 Empirical modeling

To provide an intuition for our empirical strategy, first assume that the latent factors — health status and household production — are observed and that the model (3) is linear. For a sample of n individuals observed in T time periods, the regression model is subsequently given as

$$s_{itj} = \gamma_j \eta_{itj} + \lambda_{j1} h_{ij} + \lambda_{j2} h_{ij'} + \omega_{j1} h^p_{ij} + \omega_{j2} h^p_{ij'} + \delta_j h^p_{ij} \times \eta_{itj} +$$

$$\phi_j C_{sj} + \zeta_j \mu_{ij} + \varepsilon_{itj},\ i = 1, ..., n,\ t = 1, ..., T,\ j' \neq j \tag{4}$$

where s_{itj} and η_{itj} are days of sickness absence and health status for individual i of gender j at time period t. By model assumption, ε_{itj} is a regression error; that is, i.i.d. and independent of the included covariates. Moreover, let c_i be a household indicator variable that takes the value one if an individual belongs to a household and zero otherwise. We assume (i) $\mathrm{E}(h^p_{im}|c_i = 1) = 0$, (ii) $\mathrm{E}(h^p_{im}|c_i = 0) = \mathrm{E}(h^p_{if}|c_i = 0)$ and (iii) $\mathrm{E}(h^p_{if}|c_i = 1) - \mathrm{E}(h^p_{if}|c_i = 0) > 0$. That is, we assume (i) that men with children specialize in market work and perform no home production on average, (ii) that men and women without children perform equal amounts of home production on average and (iii) that women with children on average perform strictly more home production than women and men without children. Finally, define

$$\delta^\#_f = \delta_f [\mathrm{E}(h^p_{if}|c_i = 1) - \mathrm{E}(h^p_{if}|c_i = 0)].$$

Hence, $\delta^\#_f$ measures the change in sickness absence due to a change in health stemming from the greater household production $\mathrm{E}(h^p_{if}|c_i = 1) - \mathrm{E}(h^p_{if}|c_i = 0)$.

Assume that labor supply and household production are fixed between two time periods, $t = (0, 1)$, but where individual health is poorer in the second period, i.e. $\eta_{i1j} - \eta_{i0j} = \Delta_{ij} < 0$. The change in sickness absence over the two time periods is subsequently given by

$$s_{i1j} - s_{i0j} = \gamma_j \Delta_{ij} + \delta^\#_f c_i F_i \times \Delta_{if} + \varepsilon_{i1j} - \varepsilon_{i0j},\ i = 1, ..., n_j,$$

where $n_j = n_{j1} + n_{j0}$ is the the total number of females $(j = f)$ and males $(j = m)$ aggregated over household status, $c_i = (0, 1)$.

While individual health status is unobserved, we do observe an indicator of a change in health, a hospital admission. It is useful in this context to think of the hospital admission in the following way: An individual decides to visit a hospital if his or her health falls below a certain threshold. If this threshold is different for men and women, we can write the data generating process for the hospital admission as

$$H_i = \mathbf{1}(\eta_i < \tau + \tau_F F_i),$$

where $\mathbf{1}(\cdot)$ is the indicator function (returning one if the expression inside the parenthesis is true and zero otherwise). Thus, an admission is observed if health is below a certain threshold τ for men and $\tau + \tau_F$ for women. Suppose that $\tau_F > 0$ and that average health in the time period before the visit is the same for men and women; admitted women will subsequently on average have better health than admitted men.[8]

Let $\overline{\eta}_{jt}$ be mean health of men and women, $j = (m, f)$, at time t. Then Δ_j is the mean change in health of men and women between time periods zero and one. Suppose (i) that the health of men and women is the same before the hospital admission and (ii) that men and women have the same health thresholds for visiting a hospital (i.e. $\tau_F = 0$). Under these assumptions, the health changes of hospitalized men and women will on average be equal; that is, $\Delta_m = \Delta_f = \Delta$. This means that we can identify the gender behavioral effects from the health change by estimating the following model with OLS

$$s_{it} = b_0 + b_1 F_i + \gamma_m^* H_{it} + \kappa^*(F_i \times H_{it}) + \alpha X_i + \varepsilon_{it}, \tag{5}$$

where $t = -l, \ldots, 0, \ldots, l$, $\gamma_m^* = \gamma_m \Delta > 0$ and $\kappa^* = \kappa\Delta$. Here H_{it} is a step function taking the value one for all time periods after an individual is registered for a hospital admission occurring at $t = 0$, and zero otherwise (i.e. $H_{it} = 1$ for $t > 0$ and 0 otherwise). X_i is a vector of control variables and b_0 and b_1 are the mean parameters by gender derived from the labor supply function (4). For details on the derivation see Appendix B.

This model setup makes clear that the "differences-in-differences" specification in (5) eliminates any differences in levels of sickness absence across the groups. Hence, time-invariant differences in health, economic incentives, working conditions and other

[8] We assume that doctors cannot perfectly screen the health of individuals and make correct decisions whether the visit calls for an admission or not. The implication is that if women make more visits than men they also have more admissions.

potentially confounding variables are controlled for by design. However, if men and women on average have different health thresholds when they decide to visit a hospital, $\tau_F \neq 0$ and hence $\Delta_m \neq \Delta_f$, i.e. the health shock causing an admittance will on average differ between men and women. Under the alternative hypothesis, women care more about their health than men (by e.g. investing more in their health). This implies that the (unobserved) health shocks causing the hospital admissions will on average be less severe for women than for men (i.e. $\tau_F > 0$). We show in Appendix B that under these assumptions the OLS estimator of κ^* based on equation (5) will be biased downwards. Thus, if the OLS estimate of κ^* is positive this provides a lower bound of the female/men difference in preventive behavior.

Even if the hospital admissions can be considered as objective measures of health shocks, they are also in most cases an ex-post measure of any health change observed by the individual (unless the reason for the hospital admission is due to an accident or an unexpected disease). The implication is that behavioral changes affecting sickness absence could for many health conditions occur before the actual hospital admission. In order to address this problem we have performed sensitivity analyses where we have estimated model (5) but lag the health shock one and two years. That is, we assume that the individual may observe his or her health change one and two years before we observe it in the data.

Theoretically, i.e. under the alternative hypothesis, hospital admissions for women are caused by less severe health shocks than for men and as a consequence our estimated behavior effect is biased downwards. Empirically, however, the relationship between health and gender is an open question. Thus, in order to address the concern of whether women are affected by worse health shocks than men, we evaluate the post-admission health of men and women by estimating Cox hazard regression models

$$\Pr(\text{exit}_{it} = 1 | \text{exit}_{it-1} = 0) = \lambda_0(t) \exp(\delta_1 F_i + \alpha X_i), \quad t > 0. \tag{6}$$

Here $\lambda_0(t)$ is the baseline hazard (i.e. the hazard rate for men) and δ_1 and α are parameters to be estimated. Specifically, we study the exits to death and to a new hospital admission.

A positive estimate of δ_1 for both transitions provides evidence of relatively poorer post-admission health of women compared to men.

Next, given that a gender difference in behavior in the estimation of (5) is found, we seek to further investigate whether this difference can be explained by women having household responsibilities. To this end, we can estimate the same model as in (5) above by simply substituting the female dummy for the child dummy. We show in Appendix B that when a hospital admission is being considered, if the health of women with children is better than the health of women without children (as would be the case under the alternative hypothesis), the OLS estimator will be biased downwards. Note, however, that for this analysis we can also adjust for this potential bias by including the subsample of men in the estimation. To implement this procedure we estimate a triple differences model by additionally including an indicator variable for having children along with the three first level interactions of gender, the health shock and having children. Formally, we estimate (OLS)

$$\begin{aligned} s_{it} = & \, b_0 + b_1 F_i + b_2 c_i + b_3 F_i c_i + \gamma^*_{m1} H_{it} + \gamma^*_{m0} H_{it} c_i \\ & + \gamma^*_f F_i H_{it} + \delta^*_f F_i H_{it} c_i + \varepsilon_{it}, \end{aligned} \tag{7}$$

where $\delta^*_f = \delta^\#_{f1} \Delta_f$ and γ^*_{mc} is the response for males with ($c = 1$) and without children ($c = 0$). The OLS estimator of δ^*_f is consistent under the assumption that the difference in the health threshold for being admitted to a hospital for men with or without children is greater than the difference in threshold for women with or without children (for details see Appendix B). We find support for the assumption from the gender mortality patterns displayed in *Figure 5.4*.

Finally, we also estimate Cox-regression models with exits to mortality and re-admission in which we include the indicator variable for children and its interaction with gender, i.e.

$$\Pr(\text{exit}_{it} = 1 | \text{exit}_{it-1} = 0) = \lambda_0(t) \exp(\delta_1 F_i + \delta_2 c_i + \delta_3 F_i c_i + \alpha X_i), \quad t > 0. \tag{8}$$

Here $\lambda_0(t)$ is the baseline hazard (i.e. the hazard rate for men without children). A negative estimate of δ_3 suggests that women with children compared to those without have better health. In this model we control for potential health differences between women with or without children using men with or without children as the counterfactual "effect" of having children. As we cannot control for differences in pre-hospitalization health trends for men and women without children, the estimation of the parameter of interest is subject to a stronger identifying assumption here than in the estimation of the effects on sickness absence.

3.3 How should the results be interpreted in terms of the hypotheses?

If we find that women both with and without children do not increase their absenteeism relative to men after the hospital admission, we will refute both the gender and household preference difference hypotheses, at least for our specific sample of individuals. On the other hand, if we find that women increase their sickness absence more while not having higher mortality rates than men, this would suggest that women take more preventive action than men. Moreover, if we find that these differences in sickness absence can be primarily explained by women with children, the health investment hypothesis of Paringer (1983) is subsequently supported by our data if these women also have only slightly better post-admission health in comparison to women without children. In contrast, if the higher absence rates of women with children are accompanied by higher mortality and re-admission rates for these women compared to women without children, we would reject the health investment hypothesis in favor of the double burden hypothesis (see Bratberg *et al.* (2002)).

4 Data

Our empirical analysis exploits micro data originating from administrative registers. The data, collected and maintained by Statistics Sweden, covers the entire Swedish population between 16 and 65 during the period 1987-2000, and individuals aged 16 to 74 between 2001-2010. It contains annual information on a wide range of socioeconomic variables.

Information on hospital admissions was provided by the National Board of Health and Welfare and covers all inpatient medical contacts at public hospitals from 1987 through

1996. This is no major restriction since virtually all medical care in Sweden at the time was performed by public agents. From 1997 onwards the register also includes privately operated health care. In order for an individual to be registered with a health impairment (s)he must have been admitted to a hospital. As a general rule, this means that (s)he has to have spent the night at a hospital. However, starting in 2002 the registers also cover outpatient medical contacts in specialized care.

An important feature of the data is that it contains the exact cause for each admission and death. The diagnoses, made by physicians, are classified according to the World Health Organization's International Statistical Classification of Diseases and Related Health Problems (ICD-10). ICD-10 is a seven digit coding of diseases and signs, symptoms, abnormal findings, complaints, and external causes of injury or diseases. We include all diagnosis categories in the analysis with the exception of diagnoses related to pregnancy. In order to account for potential gender heterogeneity in type of illness, we include controls for disease category in the regressions. Moreover, we also estimate separate models for the four most common groups of diseases in our data; ischemic heart diseases, musculoskeletal diseases, cancer and mental health problems.

We use the annual number of days an individual received sickness benefits as our primary dependent variable in the regressions. Information on sickness benefits was obtained from the Swedish Social Insurance Agency (SSIA), covering all individual spells of publicly paid sick leave in Sweden. Since sickness insurance is compulsory, this does not restrict our analysis with regard to the population of interest. However, since 1992, sickness spells under 14 days are no longer registered due to the introduction of an employer contribution period. Thus, in effect, we will estimate the effects of sickness spells longer than two weeks in response to the health shock. Hence, this restriction excludes absences due to minor health problems such as common colds. We do not believe this data limitation causes any problems for the analysis of the behavioral effects that we seek to identify. In fact, the restriction might even be advantageous as it may reduce variations in sickness absence between e.g. labor market sectors.

Nevertheless, there are two other potential shortcomings with the definition of sickness absence we use. First, all workers (employed and unemployed) are, in addition to the sickness insurance, also covered by a public disability insurance. These two insurances are intimately related; both are administrated by the SSIA and, as the level of compensation in the sickness insurance is higher than in the disability insurance, disabled individuals would normally prefer to be on sickness rather than on disability benefits. Hence, individuals admitted to disability insurance schemes are therefore likely to have a prehistory of sickness absence. The implication for our analysis is that being on a disability insurance scheme is also a likely outcome of the health change. Moreover, being on a disability insurance scheme may also be a relevant health investment decision for the same reason as sickness absence. To investigate this possibility we have performed sensitivity analyses where we include individuals on disability insurance schemes in the sickness absence outcome definition. Second, group differences in mortality rates could potentially bias the results as the hypothetical sickness absence level of deceased individuals is unobserved. In order to investigate this potential source of bias, we also include deceased individuals and let their absence be zero for all the subsequent years after death.

We take a random sample of 40 % of all employed[9] individuals between 20 and 50 in 1992 who were observed to have had an in-hospital care record at some point during the observation period 1993-2004. We restrict the sample to employed individuals for all the years prior to the hospital admission but, for obvious reasons, relax it thereafter.

Since most individuals have children at some point in their lives, defining a household as a unit of observation with at least one child would make household a time-varying variable. Hence, given this definition, comparing individuals with and without children at the time of the health shock would imply that household estimates on sickness absence would also reflect life cycle variation in factors such as health and sickness absence since the comparison individuals are likely to have a household at a later stage. Moreover, the health shock might also endogenously affect the probability of having a child for some individuals. In order to circumvent problems relating to a time-varying definition of households, we further restrict our sample to only include men and women who were between 40 and 45 at the time of the hospital admission and define individuals with at least one child at the age of 40 as having a household. The advantage of this definition is

[9] An individual is defined as employed if they had earnings in the tax register in November 1992.

that most individuals have completed their fertility by the age of 40 and, hence, few will change family status after their hospitalization. However, it may also potentially affect the external validity of the results from the analysis.

With the above restrictions our final sample consists of roughly 63,000 employed men and women with at least one hospital admission between 40 and 45.[10] To get an idea of how much the sample restrictions limit the inferences we can draw to the total population of working individuals, we construct a representative sample of active individuals with the same age distribution during the same period, but without a record of a hospital admission.[11]

Table A.3 and *Figure A.1–Figure A.3* in Appendix A present descriptive statistics of our analysis sample and the age-matched comparison sample. While the slightly higher wages and lower sickness absence for the comparison sample provides an indication that our analysis sample suffers poorer health, the differences are still surprisingly small. In particular, from *Figure A.1–Figure A.3* it is evident that the average differences in sickness absence and income variables between the samples are small and have similar trends in age categories prior to the hospitalizations. Based on these descriptive statistics we conclude that the analysis sample is roughly similar to the comparison sample with regard to the key variables health and income.

5 Results

This section begins with a presentation of the gender analysis followed by the household analysis. The section concludes with a discussion on the results from a number of robustness analyses.

[10]The original population of all labor market active individuals between 20-50 in 1993 amounted to 2,587,580 individuals. After taking a random sample of 40% (leaving 1,035,032 individuals) and conditioning on individuals having been hospitalized between 1993-2004 (leaving 470,587 individuals), aged between 40 and 45 at the time of hospitalization (leaving 75,880 individuals) and finally removing individuals with labor market interruptions in the years before the hospital admission we are left with 63,599 individuals in our analysis sample. See *Table A.2* in Appendix A for descriptive statistics of the sample and the included variables.

[11]Specifically, each sampled individual from our analysis sample (i.e. with a hospital admission record) is matched with a person of the same age without a hospital admission record in the age range 40-45.

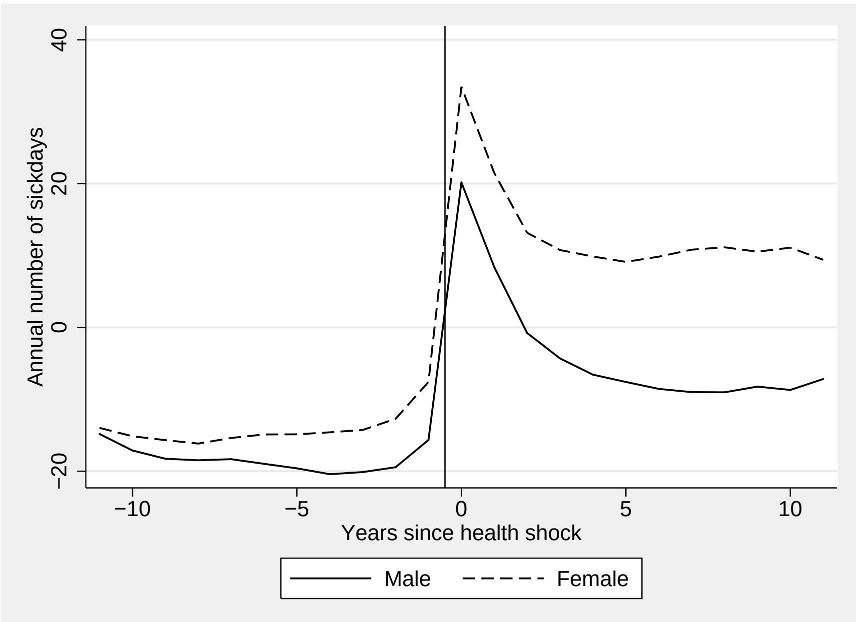
5.1 Gender differences in sickness absence and health

We first present the results on sickness absence followed by the results from the post-admission health outcomes, i.e. the probabilities of mortality and re-hospitalization.

5.1.1 Sickness absence

Before turning to the results from our estimations we first present some graphical evidence. *Figure 5.1* plots the average number of sick days for men (solid line) and women (dashed line) by years from the hospital admission. The figure clearly shows that women have on average a greater number of sick days both before – but in particular after – the hospital admission compared to men. This observation is indicative of men and women exerting different absence behavior after the hospitalization spell. From the figure we can also see that the gender difference in sickness absence increases slightly before the admission. We would expect this pattern if women were acting in a more preventative way than men, since the individuals in our sample most likely observed their health change before we observed it from the hospital admission. We discuss this potential problem below.

Figure 5.1: The average number of annual days of sickness absence for men and women by years after a hospital admission



NOTE.—The table is constructed by plotting the residuals from an ordinary least squares regression of year fixed effects on days of sickness absence. Days of sickness are defined as the number of days on sickness absence during one year. The male (female) average number of sick days is measured by the solid (dashed) line. The vertical line indicates the hospital admission.

Table 5.1 presents the estimated κ^* from equation (5) using the same observations making up *Figure 5.1*. Hence, the parameter we estimate is the annual effect on gender difference in sickness absence over in total twelve years after the hospitalization. The estimated results confirm the pattern displayed in *Figure 5.1.* Our preferred specification, given in column (3) of *Table 5.1*, shows that women have on average an additional twelve days of sickness absence after their spell in hospital compared to men.[12] Note that the inclusion of control variables only marginally influences the estimated parameters. To the extent that the estimates change, however, the inclusion of fixed year and age effects and other control variables slightly increases the estimated gender difference in sickness absence. Given that the inclusion of control variables captures some of the gender differences in health, this pattern indicates that women on average have better pre-admission health than men (see Appendix B for an informal proof of this proposition).

Thus far, our results suggest that women in general act in a more preventative way than men by utilizing sickness insurance more when they experience an adverse change in health. However, if our assumption — that the health changes of men and women in our sample are comparable — is invalid, these estimated effects may potentially be the result of gender differences in the type of illnesses affecting the individuals. While the inclusion of diagnosis and labor market sector fix effects left our results qualitatively unchanged, heterogeneous effects may still exist across these categories and sectors. In particular, *Table A.4* and *Table A.5* show that men and women in our sample are generally affected by different types of illnesses and work in different sectors. The range is substantial; from a female share of .84 for cancer to only .37 for heart diseases and from a share of .87 in the health sector to .09 in the construction sector.

Figure 5.2 plots the sickness absence of men and women before and after the observed hospital admission for the four most common diseases in our data; cancer, heart, mental and musculoskeletal diseases. Interestingly, the pre-admission trends are basically parallel in all panels, suggesting that the aggregate gender difference in the pre-admission sickness absence trend is driven by gender differences in diagnosis category at least to

[12]We have also performed analyses where we have estimated the effects 2, 4, 6, 8 and 10 years after the hospital admission.

Table 5.1: The estimated relative effect on days on sickness absence from a hospital admission for women relative to men

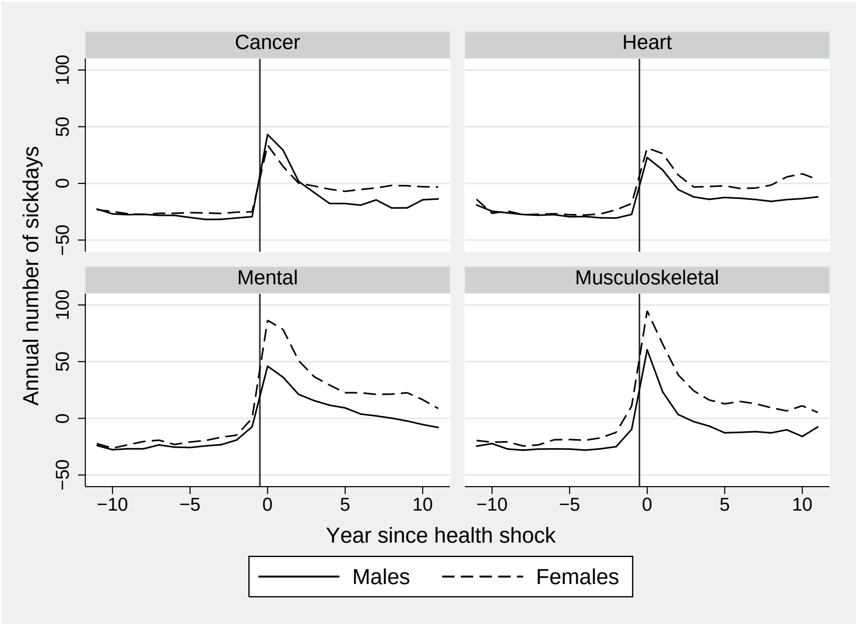
	(1)	(2)	(3)
Full sample	10.980***	11.401***	12.582***
	(0.319)	(0.318)	(0.317)
Diagnosis type			
Heart	7.727***	8.440***	9.328***
	(1.257)	(1.259)	(1.261)
Cancer	2.188	2.941**	4.019***
	(1.480)	(1.479)	(1.470)
Mental	17.799***	18.558***	21.251***
	(1.851)	(1.849)	(1.867)
Musculoskeletal	15.352***	17.153***	17.862***
	(1.438)	(1.431)	(1.435)
Industry Sector			
Manufacturing	9.798***	10.316***	11.100***
	(0.897)	(0.896)	(0.890)
Public	14.110***	14.513***	15.208***
	(1.097)	(1.097)	(1.095)
Education	13.667***	12.899***	12.672***
	(1.137)	(1.133)	(1.131)
Health	11.607***	12.045***	12.271***
	(0.966)	(0.965)	(0.959)
Year/Age FE	No	Yes	Yes
Additional controls	No	No	Yes

NOTE.—The table reports the estimated parameter (standard error) of the female and hospital admission interaction effects for different samples. In the estimation we control for the main effect. Estimation is performed with ordinary least squares. Standard errors are estimated using a robust covariance matrix.*p<0.1, **p<0.05 ***p<0.01. The second column include year and age fixed effects and the last column also include 15 industry and 19 diagnosis category fixed effects along with additional controls for annual income and dummy variables for high earner and high education. See *Table A.1* for detailed variable definitions.

some extent. The sickness absence of males is higher in the first year after admission for cancer patients but falls below that of women in subsequent years. For the other categories the post-admission pattern closely follows the aggregate results.

Results from a separate estimation of model (5) for each diagnosis category and four different labor market sectors (manufacturing, public administration, education and health) are displayed in rows 2-9 of *Table 5.1*. What is noteworthy is the consistent pattern of a relative increase in female absenteeism irrespective of category and/or the inclusion of controls. Moreover, adding control variables increases the difference slightly for all sub-analyses, which provides evidence that women on average have better pre-admission health also within each diagnosis and labor market sector category (see Appendix B for details).

Figure 5.2: The average number of annual days of sickness absence by gender and years after a hospital admission, by diagnosis type



NOTE.—The figures are constructed by plotting the residuals from an ordinary least squares regression of year fixed effects on days of sickness absence. Days of sickness are defined as the number of days on sickness absence during one year. The male (female) average number of sick days is measured by the solid (dashed) line. Each panel pertains to a specific diagnosis type. The vertical line indicates the hospital admission.

Interestingly, from rows 2-5 of *Table 5.1* we find quite a large variation in the estimated gender effects on sickness absence across different diagnoses. In particular, a woman with a cancer diagnosis has roughly four more days of sickness absence after the admission than a man with a cancer diagnosis. In contrast, a woman with a mental disease has on average over twenty days more of sickness absence than a man with the same diagnosed condition. For the more vague diagnoses it is more likely that the magnitude of sickness absence is more affected by a patient's own interpretation of the illness and less by a physician's assessment of the illness (see e.g. (Englund, 2001)). Hence, this results provides further evidence of a gender behavior effect in relation to the health shock.

Finally, the results from a separate estimation of the four different labor market sectors are displayed in rows 6-9. These results are remarkably consistent across sectors which provide strong evidence that the gender differences in the response are not driven by the gender segregated labor market in Sweden.

5.1.2 Mortality and hospital admissions

The results in the previous section showed that there was a relative increase in the sickness absence of women after a spell in hospital compared to men. However, it is possible that men and women are affected in different ways by similar health shocks, even within diagnostic categories, and so the question remains whether these differences in sickness absence are driven by poorer post-admission health among women, rather than gender differences in health behavior.

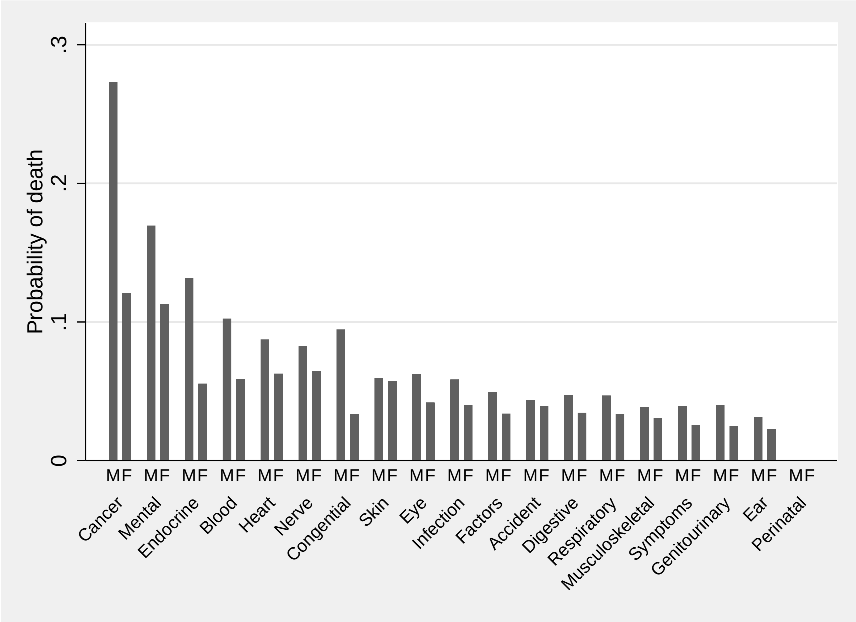
As a starting point we present sample statistics of the fraction of men and women who died within three years after being admitted to hospital for each diagnosis category in *Figure 5.3*. *Figure 5.4* further categorizes these three-year mortality fractions by family status. The resulting pattern is unambiguous; men have a higher mortality rate for all diagnosis categories. In particular, the mortality rate for males is more than twice as high as that of women for the subsample of cancer diagnosis (.27 compared to .12). Moreover, the difference in cancer mortality is extreme for men and women without households (.41 compared to .13). Since cancer is a disease for which the mortality risk is strongly correlated with the time of detection (see e.g. Levin *et al.* (2008), Brett (1969) and UK trial of early detection of breast cancer group (1988)), this massive gender difference in cancer mortality provides a hint of how a more preventative health behavior among men could enhance their longevity. This difference is also very likely to be a consequence of the public screening program for breast cancer in Sweden which has been ongoing since 1986.[13] Finally, the large gender difference in the cancer mortality rate may also explain some of the relatively small estimated differences in sickness absence between men and women diagnosed with cancer in the previous section.

Next, *Table 5.2* presents estimation results from the Cox regression model (6) with exits to death and a second hospital admission.[14] The first column (1) displays the results without controls while column (2) displays the results including the full set of controls.

[13] There is an important discussion questioning the value of screening for breast cancer today see e.g. McPherson (2010). This criticism should, however, not be interpreted as criticism of the value of early detection. Today most women are aware of the risk of breast cancer and of the possibilities of self-screening which was not the case almost 30 years ago when screening was introduced.

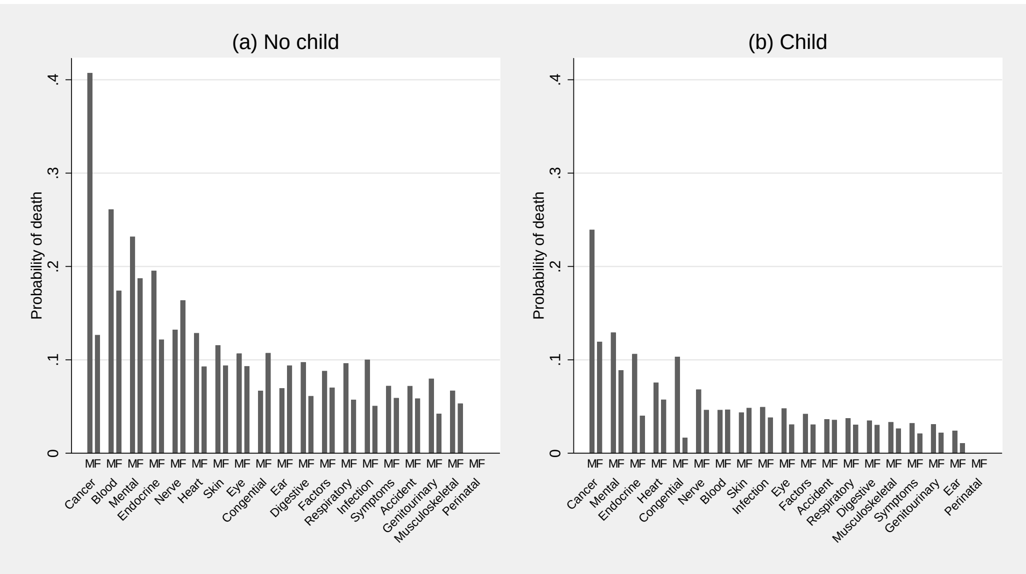
[14] We have also estimated linear probability models for the same outcomes occurring within two, three etc. years after the hospital admission. The results are qualitatively similar to the Cox regression model.

Figure 5.3: Mortality risk after a hospital admission, by gender and diagnosis type



NOTE.—Mortality is the fraction of individuals who died within three years after hospital admission. The bins of the histograms pertains to (F)emales and (M)ales for each of the 19 diagnosis categories in diagonal text under the x-axis as displayed in *Table A.4*.

Figure 5.4: Mortality risk after a hospital admission, by gender, household status and diagnosis type



NOTE.—Mortality is the fraction of individuals who died within three years after hospital admission. The bins of the histograms pertains to (F)emales and (M)ales for each of the 19 diagnosis categories in diagonal text under the x-axis as displayed in *Table A.4*.

As in *Table 5.1*, the first row of *Table 5.2* shows the estimation results for the full sample while rows 2-4 (5-8) display the results when separately estimating the model for the four different diagnoses (sectors).

Table 5.2: The estimated post-admission relative mortality and re-admission risks for women relative to men

	Death		Admission#2	
	(1)	(2)	(1)	(2)
Full sample	-0.211***	-0.326***	0.049***	0.003
	(0.046)	(0.048)	(0.012)	(0.013)
Diagnosis type				
Heart	-0.590***	-0.691***	-0.042	-0.097**
	(0.193)	(0.197)	(0.043)	(0.045)
Cancer	-0.951***	-1.099***	-0.293***	-0.380***
	(0.090)	(0.094)	(0.051)	(0.054)
Mental	-0.521***	-0.498***	0.022	0.011
	(0.138)	(0.142)	(0.045)	(0.046)
Musculoskeletal	-0.264	-0.372	0.273***	0.220***
	(0.221)	(0.228)	(0.043)	(0.045)
Industry Sector				
Manufacturing	-0.155	-0.295**	0.071**	0.019
	(0.124)	(0.126)	(0.031)	(0.032)
Public	-0.337*	-0.443**	0.061	-0.005
	(0.188)	(0.196)	(0.046)	(0.049)
Education	-0.031	-0.109	0.064	0.034
	(0.185)	(0.189)	(0.048)	(0.049)
Health	-0.205	-0.272*	-0.002	-0.097**
	(0.157)	(0.164)	(0.037)	(0.039)
Controls	No	Yes	No	Yes

NOTE.—The table reports the estimated coefficient (standard error) on the female/male relative hazard to death and a second hospital admission. Estimation is performed under the asumption of a Cox proportional hazards model taking use of an exact maximum likelihood estimator. *p<0.1, **p<0.05 ***p<0.01. The regressions in the second column of each dependent variable control for year and age fixed effects and additional controls for annual income and dummy variables for high earner and high education. See *Table A.1* for detailed variable definitions.

From the first row of column (1) we find that women have an estimated lower risk of death but a higher risk of a second admission — both estimates being highly significant — after the hospital admission. However, when adding controls to the models, the negative effect on mortality increases while the effect on a second admission becomes insignificant and close to zero. Even though we include a rich set of control variables, it is still likely that we have missed some important health variables in the estimation. The implication of this is that, if anything, the estimates in columns (2) should still both be biased towards zero. Thus, all in all we interpret these results as establishing that the women in our sample have on average better health — not only before — but also after the hospital admission than the men in our sample.

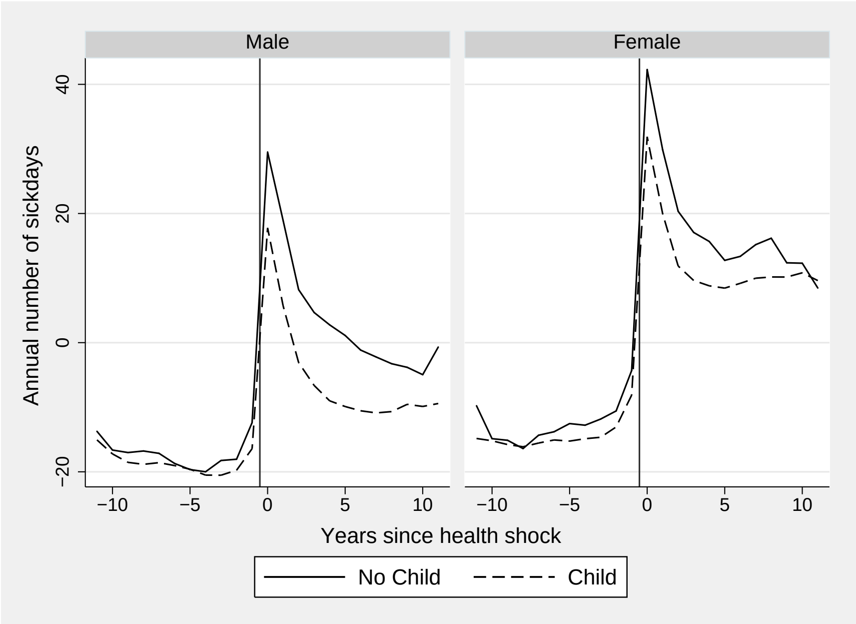
The results conditional on having a specific diagnosis (shown in rows 2-5) are in general consistent with the aggregate results for the entire sample. The negative effect on mortality is, as expected, greater and, adding control variables increases (decreases) the magnitude of the negative (positive) estimates. Women have an estimated lower, albeit insignificant, risk of mortality for musculoskeletal diseases but have also a significantly greater risk of a hospital re-admission (also when including control variables). Since musculoskeletal diseases are, in general, more of a symptom diagnosis than cancer, for example, we interpret the increased risk of re-admission for this diagnosis rather as the effect of a greater degree of preventive behavior among women than a consequence of poorer post-admission health. Finally, the results from the within-industry regressions (rows 6-9) are generally estimated with lower precision but the resulting point estimates are much in line with the aggregate results.

5.2 Household differences in sickness absence and health

Before discussing the estimation results we provide some initial graphical evidence. In *Figure 5.5* the average number of sick days for men (left panel) and women (right panel) with (dashed line) and without children (solid line) is illustrated. The figure shows that individuals with children have on average less absenteeism both before and after the admission. This pattern was expected given the descriptive sample statistics in *Table A.3*. The figure further reveals that while the increase in sickness absence after the admission is greater for individuals without children, it is less pronounced for women. Hence, this descriptive analysis provides some preliminary evidence in support of both the health investment hypothesis of Paringer (1983) and the double burden hypothesis of Bratberg *et al.* (2002); i.e., women with children have relatively more days of absence than women without children following a change in individual health.

The estimation results of two different regression models using the same observations making up *Figure 5.5* are displayed in panel A of *Table 5.3*. The estimates presented in columns (1) through (6) are the results where we have estimated regression models for each panel of *Figure 5.5* separately, while the results in column (7) refer to the results when we have estimated equation (7), in other words we have subtracted the right panel from the left. We show in Appendix B that if the health of women/men with children

Figure 5.5: The average number of annual days of sickness absence by children and years after a hospital admission, by gender



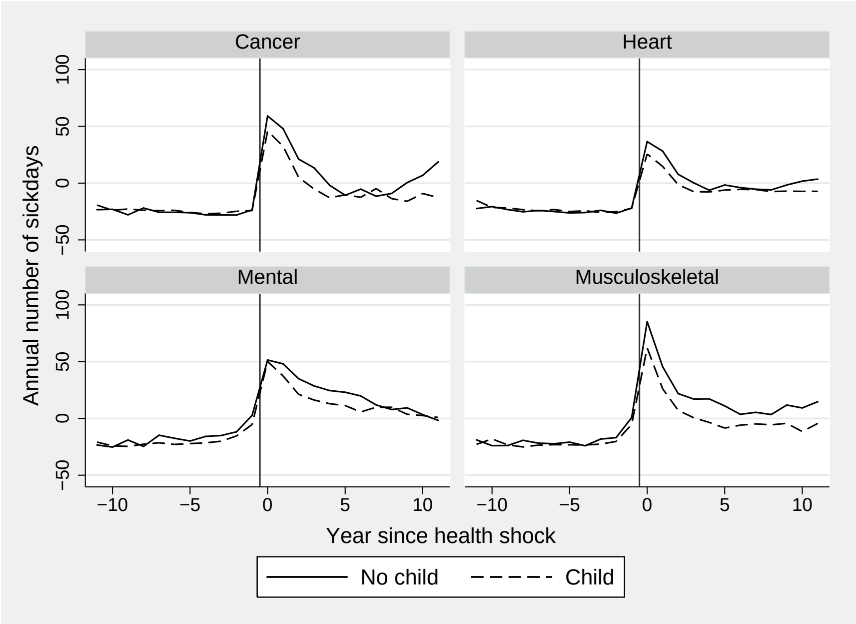
NOTE.—The figures are constructed by plotting the residuals from an ordinary least squares regression of year fixed effects on days of sickness absence. Days of sickness are defined as the number of days on sickness absence during one year. Each panel pertains to a specific diagnosis. The dashed (solid) line pertains to the average number of days of sickness absence with (without) children. The vertical line indicates the hospital admission.

requiring a hospital admission is better than the health of women/men without children, the estimated effects in columns (1) through (6) are biased downwards.

In line with the pattern seen in *Figure 5.5*, *Table 5.3* shows that the response from the admission is lower for individuals with children, irrespective of gender. However, while men with children increase their absence between five to seven days less in comparison to men without children, the corresponding difference for females is smaller: only between one and three days. The estimated coefficients in the regressions are, as expected, lower when we add control variables (for details see Appendix B). The bias thus moves in the same direction as in the gender comparison, suggesting that the health before the admission is better for those with children than that of those without children. From our main specification, displayed in column (7), we can see that, in contrast to women without children, women with children increase their absenteeism by on average five additional days each year after the admission.

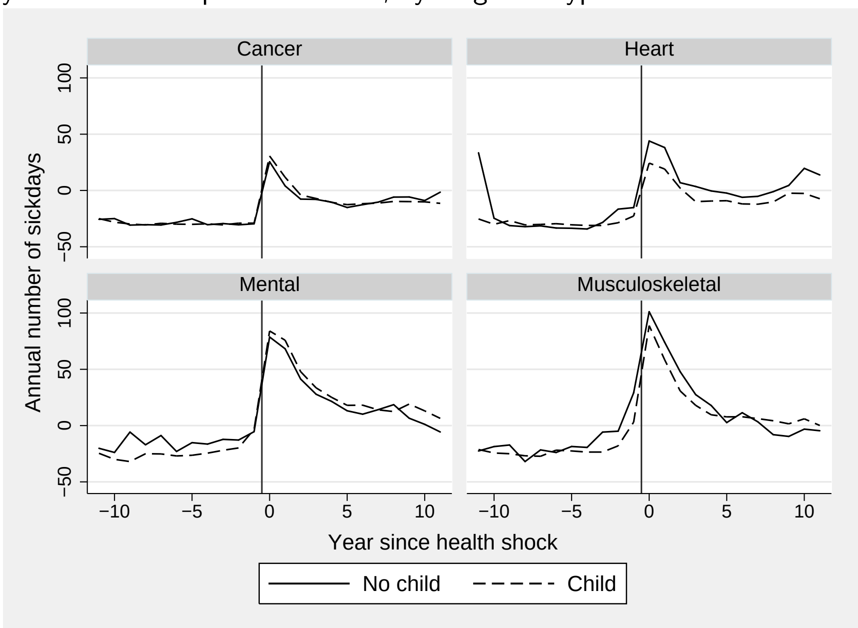
Turning to the sub-analyses by diagnosis category, the results from the estimation conditional on each of the four selected diagnosis categories are displayed in rows 2 to

Figure 5.6: The average number of annual days of sickness absence for men by family status and years after a hospital admission, by diagnosis type



NOTE.—The figures are constructed by plotting the residuals from an ordinary least squares regression of year fixed effects on days of sickness absence. Days of sickness are defined as the number of days on sickness absence during one year. Each panel pertains to a specific diagnosis. The dashed (solid) line pertains to the average number of days of sickness absence with (without) children. The vertical line indicates the hospital admission.

Figure 5.7: The average number of annual days of sickness absence for women by family status and years after a hospital admission, by diagnosis type



NOTE.—The figures are constructed by plotting the residuals from an ordinary least squares regression of year fixed effects on days of sickness absence. Days of sickness are defined as the number of days on sickness absence during one year. Each panel pertains to a specific diagnosis. The dashed (solid) line pertains to the average number of days of sickness absence with (without) children. The vertical line indicates the hospital admission.

5 in *Table 5.3*. *Figure 5.6-Figure 5.7* also plot the average number of days of sickness absence over time for each diagnosis category by family status, separately for men and women. The pattern for men is in line with the aggregate results while the pattern is more heterogeneous for women. For cancer or a mental diagnosis, women with children increase their absence more than women without children. This is also the case with regard to the long-run effect for women with a musculoskeletal diagnosis. The results from the OLS estimation, displayed in rows 2-5 of the table, correspond closely to the pattern displayed in the figures. The results are also consistent with the estimates from the aggregate results in the first row in *Table 5.3* with the exception of admissions with a heart diagnosis where the estimated effect cannot exclude zero at any conventional risk level. It is interesting to note that the differences in sickness absence again increase with the vagueness of the diagnosis. This suggests that any potential "investment effect" is greatest in a diagnosis where individual freedom with regard to deciding whether or not to be absent from work is greatest, i.e. for mental and musculoskeletal diseases.

Table 5.3: The estimated relative effect on days on sickness absence from a hospital admission for individuals with households (children)

	Males			Females			Both genders
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. All diagnoses included							
Full sample	-7.497***	-7.641***	-5.163***	-2.754***	-2.859***	-1.020	4.622***
	(0.541)	(0.540)	(0.542)	(0.729)	(0.728)	(0.727)	(0.905)
Diagnosis type							
Heart	-5.989***	-6.087***	-4.638***	-5.917*	-6.049*	-6.045*	-0.694
	(1.712)	(1.712)	(1.732)	(3.202)	(3.177)	(3.195)	(3.636)
Cancer	-12.274***	-12.752***	-10.335***	1.669	1.675	2.810*	13.511***
	(3.753)	(3.732)	(3.736)	(1.520)	(1.516)	(1.519)	(4.035)
Mental	-1.289	-3.322	-0.657	13.683***	10.701***	9.827***	14.362***
	(2.451)	(2.453)	(2.493)	(3.542)	(3.552)	(3.626)	(4.315)
Musculoskeletal	-11.917***	-12.775***	-10.383***	11.703***	9.777***	10.049***	23.678***
	(2.424)	(2.412)	(2.401)	(3.695)	(3.652)	(3.664)	(4.382)
Panel B. Excluding Cancer diagnoses							
Full sample	-7.373***	-7.496***	-4.888***	-3.345***	-3.476***	-1.507*	4.101***
	(0.548)	(0.547)	(0.546)	(0.830)	(0.828)	(0.826)	(0.987)
Industry Sector							
Manufacturing	-8.284***	-8.359***	-8.129***	-5.242**	-4.828**	-5.041**	2.621
	(1.004)	(0.999)	(0.991)	(2.324)	(2.321)	(2.309)	(2.502)
Public	-8.332***	-8.876***	-7.322***	3.142	3.459	4.421	12.138***
	(1.980)	(1.994)	(1.974)	(2.881)	(2.880)	(2.875)	(3.473)
Education	-6.475**	-6.195**	-4.961*	-2.194	-2.675	-4.103	1.484
	(2.588)	(2.580)	(2.599)	(2.859)	(2.856)	(2.862)	(3.868)
Health	-6.315***	-6.603***	-7.054***	-10.112***	-10.149***	-9.333***	-2.559
	(2.392)	(2.385)	(2.374)	(1.503)	(1.499)	(1.490)	(2.808)
Year/Age FE	No	Yes	Yes	No	Yes	Yes	Yes
Additional controls	No	No	Yes	No	No	Yes	Yes

NOTE.—Columns (1)-(6) in the table reports coefficient (standard errors) of a dummy variables for having children interacted with an hospital admission on the number of days on insured sickness absence. In the estimation we control for the main effect. The last column reports the triple interaction ((child*female*family) coefficient. In the estimation we control for the main effects and their second order interactions. Standard errors are estimated using a robust covariance matrix. *p<0.1, **p<0.05 ***p<0.01. The second and fourth column include year and age fixed effects and the third, sixth and seventh column also include 15 industry and 19 diagnosis category fixed effects along with additional controls for annual income and dummy variables for high earner and high education. See *Table A.1* for detailed variable definitions.

In order to investigate the potential concern of heterogenous health shocks for women with and without children we present the results from the Cox regressions in *Table 5.4*. The preferred estimates using model specification (5) are displayed in columns (5) and (6). For the sake of completeness we also present the results from a separate estimation by gender, displayed in columns (1)-(2) and (3)-(4) for males and females respectively. Columns (1) and (3) in the first row of panel A of *Table 5.4* show a lower mortality risk for both men and women with children: the annual mortality risk is approximately 36 percent lower for women and approximately 56 percent lower for men.[15] Consequently, the results displayed in column (5), in which we include the men to capture health differences by family status of the women, we also find that the mortality risk increases with around 33 per cent for a women with household responsibilty. The same pattern can be seen for the risk of a hospital re-admission; that is, lower risks for individuals with children but a relatively lower risk reduction for women compared to men. All in all, these estimates suggest a larger negative effect on health after the admission for women with household responsibility. This would then allow us to reject the health investment hypothesis.

[15]This risk reduction is obtained as 100*(exp(coefficient) − 1).

Table 5.4: The estimated post-admission relative mortality and re-admission risks for individuals with households (children)

	Males		**Females**		**Both genders**	
	Death (1)	Admission#2 (2)	Death (3)	Admission#2 (4)	Death (5)	Admission#2 (6)
Panel A. All diagnoses included						
Full sample	-0.894***	-0.188***	-0.484***	-0.123***	0.416***	0.066**
	(0.065)	(0.021)	(0.079)	(0.022)	(0.102)	(0.030)
Diagnosis type						
Heart	-0.673***	-0.115*	-0.161	-0.121	0.555	-0.006
	(0.207)	(0.062)	(0.424)	(0.091)	(0.468)	(0.110)
Cancer	-0.687***	-0.099	0.111	0.071	0.791***	0.183
	(0.162)	(0.112)	(0.130)	(0.056)	(0.208)	(0.125)
Mental	-0.587***	-0.296***	-0.432*	-0.215***	0.102	0.089
	(0.166)	(0.062)	(0.238)	(0.075)	(0.286)	(0.096)
Musculoskeletal	-1.026***	-0.181**	-0.752*	-0.239***	0.367	-0.060
	(0.292)	(0.079)	(0.384)	(0.078)	(0.475)	(0.110)
Panel B. Excluding Cancer diagnoses						
Full sample	-0.953***	-0.191***	-0.779***	-0.163***	0.165	0.028
	(0.072)	(0.021)	(0.101)	(0.024)	(0.123)	(0.032)
Industry Sector						
Manufacturing	-0.763***	-0.173***	-0.626**	-0.160**	0.019	0.001
	(0.148)	(0.041)	(0.305)	(0.071)	(0.334)	(0.082)
Public	-0.862***	-0.131	-1.126***	-0.263***	-0.287	-0.134
	(0.318)	(0.088)	(0.373)	(0.085)	(0.486)	(0.122)
Education	-0.844**	-0.172*	-0.915**	-0.244***	-0.008	-0.066
	(0.349)	(0.102)	(0.362)	(0.082)	(0.498)	(0.130)
Health	-0.814**	-0.167**	-0.440**	-0.146***	0.424	0.062
	(0.333)	(0.081)	(0.223)	(0.042)	(0.397)	(0.090)

NOTE.—Columns (1)-(6) in the table reports the estimated coefficient (standard error) on the relative (child/(no child)) hazard to death and a second hospital admission. Estimation is performed under the asumption of a Cox proportional hazards model taking use of an exact maximum likelihood estimator. Columns (1)-(2) pertain to the results for males and columns (3)-(4) to females. Columns (5)-(6) reports the estimated interaction coefficient (standard error) of gender and having children from regressing the same model but instead including both genders and controlling for the level effects of both variables. *p<0.1, **p<0.05 ***p<0.01. All regressions include controls for year and age fixed effects and additional controls for annual income and dummy variables for high earner and high education. See *Table A.1* for detailed variable definitions.

The above inference seems to be primarily driven by the subsample of cancer patients. Men and women suffer different forms of cancer; for this reason men admitted to hospital for cancer may not be valid to control for differences in unobserved health across women as for other diagnoses, especially when we observe such large differences in mortality rates for men with, as compared to without, children (see *Figure 5.4*). A potential reason for the large differences in mortality rates across the genders and for men with and without children is the ongoing (cervical and breast) public cancer screening program in Sweden which only targets women. The potential effect from screening affecting gender differences in cancer survival is clear; however, the large difference among men across family status could additionally stem from the fact that men with families are indirectly affected by the screening of their female spouses leading to an increased awareness of the benefits of early cancer detection. For these reasons, it is debatable whether our identifying assumption regarding the validity of using men as controls for the counterfactual effect of having children is valid for the case of cancer patients. The next subsection presents the results after the exclusion of individuals with a cancer diagnosis from the analysis.

5.2.1 Results from excluding cancers

Panel B of *Table 5.4* reports the results after we have excluded cancer patients from the analysis. From columns (1) and (3) we once again see an estimated lower mortality risk for individuals with children. The yearly mortality risk is approximately 61% and 54% lower for men and women respectively. When including men as controls for unobserved differences in health across family status, the estimated difference in post-admission mortality for women with household responsibility decreases sharply in magnitude and becomes insignificant (while still being fairly precisely estimated). Moreover, the same thing is true for the risk of a hospital re-admission (from columns (2), (4), and (6)). The results are also robust after performing separate analyses for the four industry sectors (see rows 2-5 of panel B). Hence, when we exclude individuals with a cancer diagnosis, we cannot find that household responsibility is related to worse post-admission health.

Next, we proceed and investigate whether the results regarding sickness absence remain unchanged by estimating the same sickness absence models as for the full sample, but now using the subsample in which cancer patients have been excluded. The estimation results are displayed in panel B of *Table 5.3*. These results are very similar to the results for the full analysis sample displayed in panel A. In particular, we find that having household responsibility increases sickness absence by on average four days after the admission. Finally, in rows 2 to 5 of panel B, we also report the results from the analysis for each of the four labor market sectors separately. Three out of four sectors report positive point estimates but due to lower precision of smaller sample sizes only the result for the public sector is statistically significant.

5.3 Robustness checks

We have performed a number of robustness checks in order to validate our main results, most of which are discussed at length in the previous sections. Here, we only provide a brief summary of the analyzes made and the reported results.

First, we re-estimated the models for sickness absence, hospital re-admission and mortality when restricting the outcome variables to only include observations less than or equal to two, four, six, eight and ten years after the hospital admission (compared to twelve years in the main specification). These results (not reported here) are very similar to our main specifications. If anything, the effects on sickness absence increase somewhat with the time window: for the gender analysis the difference in sickness absence increases from 8.9 days two years after the admission to 12.6 days ten years after the admission. The corresponding figures are 3.2 and 3.9 days for the household analysis.

Second, we re-estimated the post-admission health outcome models where we additionally controlled for the length of sickness absence and diagnosis category using stratified Cox regression models. The argument underlying this robustness check is simple: under the null hypothesis of similar gender absence behavior, the length of sickness absence and the diagnosis category are valid measures of the gravity of pre-admission health. As a consequence, we should not find any differences in mortality and re-admission once we control for these factors. Interestingly, the estimation results (not reported here) from stratified Cox regression models provide the same inference as in our main specification. In addition, the proportional hazards assumption in the Cox regression models could not be rejected for our data.[16]

In addition to these robustness checks, we have also tested the sensitivity of the results when varying the definition of the outcome variables. First, the observed measure of the health shock we use in the analysis is likely to have occurred after the actual health change observed by the individual. As a result the behavioral effect on sickness absence could in many cases have started already before the admission to hospital. Second, we have estimated the effects on sickness absence prevalence. Here this is defined as a binary variable which takes the value one if the individual received any sickness benefits during the year, and zero otherwise. Third, we have included individuals on disability benefits a given year by setting their sickness absence to 365 days. Fourth and lastly, we have retained all deceased individuals in the estimation and set their sickness absence to zero for all the post-death years.

Table 5.5 reports the results from the estimation of the same regression models as before but where we have instead lagged the hospital admission by one and two years. By lagging the admission we arrive at less precise estimates; however, the results for the gender analysis are remarkably stable while the results from the household analysis are somewhat less stable, but the general pattern remains the same. The estimated effects for the gender and household analyses using our continuous outcomes are presented in columns (1) and (3) in *Table 5.6*. The corresponding gender/household estimates for our prevalence outcomes are presented in (2) and (4) in the same table.

For ease of comparison we also report the results from the baseline model, but after we have excluded individuals with a cancer diagnosis, in the first row of the table. In this row it is clear that inference does not depend on the type of outcome measure. Turning to the second row, which reports the results when we have included disability beneficiaries in the continuous outcome, we can see that the estimated gender difference increases for both outcomes. Considering the household analysis, the inference using the prevalence outcome is unchanged while the results using our continuous outcome are no longer statistically significant. This suggests a higher probability that women without children enter

[16]The key assumption of proportionality is tested by analysing the Schoenfeld residuals following the generalization by Grambsch and Therneau (1994).

Table 5.5: Robustness checks I: Specifying different times until the onset of the health shock

	Gender differences model			Household differences model		
	Number of years until shock					
	0 years	1 years	2 years	0 years	1 years	2 years
Full sample	12.582***	11.321***	11.502***	4.101***	3.349***	4.192***
	(0.317)	(0.386)	(0.440)	(0.905)	(1.124)	(1.291)
Diagnosis code						
Heart	9.328***	7.079***	7.846***	-0.694	-4.536	-6.609
	(1.261)	(1.602)	(1.851)	(3.636)	(4.814)	(5.620)
Cancer	4.019***	4.060**	1.904	13.511***	10.160*	12.492**
	(1.470)	(1.916)	(2.220)	(4.035)	(5.227)	(6.102)
Mental	21.251***	14.856***	18.418***	14.362***	14.801**	9.081
	(1.867)	(2.470)	(2.923)	(4.315)	(5.878)	(7.235)
Musculoskeletal	17.862***	16.549***	16.975***	23.678***	18.413***	21.155***
	(1.435)	(1.814)	(2.091)	(4.382)	(5.498)	(6.441)
Industry Sector						
Manufacturing	11.100***	7.867***	6.797***	2.621	1.665	1.895
	(0.890	(1.051)	(1.189)	(2.502)	(2.857)	(3.178)
Public	15.208***	11.954***	14.491***	12.138***	9.484**	4.858
	(1.095)	(1.352)	(1.540)	(3.473)	(3.877)	(4.481)
Education	12.672***	12.962***	15.504***	1.484	5.857	6.521
	(1.131)	(1.387)	(1.544)	(3.868)	(4.439)	(3.970)
Health	12.271***	9.726***	9.733***	-2.559	-0.259	6.538*
	(0.959)	(1.173)	(1.338)	(2.808)	(3.402)	(3.970)

NOTE.—Columns (1)-(3) in the table reports the estimated parameter (standard error) of the female and hospital admission interaction effects. Columns (4)-(6) reports the triple interaction ((admisson*child*female) coefficient. In the estimations we control for the main effects and, for columns (4)-(6), their second order interactions. Each column pertains to a different hypothesized onset of the health shock measured by zero to two years before the observed hospital admission. Estimation is performed with ordinary least squares. Standard errors are estimated using a robust covariance matrix.*p<0.1, **p<0.05 ***p<0.01. The outcome in all columns is the average annual days of insured sickness absence. The first and fourth column rephrases the earlier results from s *Table 5.1* and *Table 5.3*. Each row pertains to a different sample as indicated by the row name. All specifications include year and age fixed effects, 15 industry and 19 diagnosis category fixed effects along with additional controls for annual income and dummy variables for high earner and high education. See *Table A.1* for detailed variable definitions.

the disability insurance program after their spell in hospital. In the third and fourth row respectively, we present the results for each outcome for the intention to treat analysis, i.e. where we set the absence of deceased individuals to zero. From these two rows it is clear that inferences are not affected by the inclusion of the deceased individuals.[17]

In conclusion, the results for the gender analysis are remarkably robust to all the robustness analyses we have performed in this section. The results for the household analysis are somewhat more mixed, mainly due to more imprecisely estimated parameters. Including disability beneficiaries in the measure of sickness absence reduces the magnitude of the estimated effect and is only statistically significant for the prevalence outcome.

[17] We have also estimated the same models keeping individuals with cancer with qualitatively the same results.

Table 5.6: Robustness checks II: Different outcome measures and including deceased individuals

	Gender differences model		Household differences model	
	Days of absence	Prevalence	Days of absence	Prevalence
	(1)	(2)	(3)	(4)
Baseline	12.582***	0.045***	4.101***	0.025***
	(0.336)	(0.002)	(0.987)	(0.006)
Include disability	24.743***	0.067***	1.744	0.020***
	(0.443)	(0.002)	(1.380)	(0.006)
ITT	12.754***	0.046***	3.968***	0.024***
	(0.333)	(0.002)	(0.969)	(0.006)
ITT and disability	25.005***	0.068***	1.441	0.018***
	(0.440)	(0.002)	(1.366)	(0.006)

NOTE.—Columns (1)-(2) in the table reports the estimated parameter (standard error) of the female and hospital admission interaction effects. Columns (3)-(4) reports the triple interaction ((admisson*child*female) coefficient. In the estimations we control for the main effects and, for columns (3) and (4), their second order interactions. Estimation is performed with ordinary least squares. Standard errors are estimated using a robust covariance matrix.*p<0.1, **p<0.05 ***p<0.01. The outcome measure in columns (1) and (3) are the avereage annual days of insured sickness absence and the outcome measure in columns (2) and (4) is an indicator for whether the individual was on disability or sickness benefits during the ex-post period. The baseline row rephrases the earlier results from s *Table 5.1* and *Table 5.3*. The 'include disability' rows includes disability beneficiaries (in columns (1) and (3) and individual on disability benefits are set to be work absent for 365 days). The 'ITT' row includes deceased individuals. Robust standard deviations are reported in parentheses with significance levels equal to *p<0.1, **p<0.05 ***p<0.01. All specifications include year and age fixed effects, 15 industry and 19 diagnosis category fixed effects along with additional controls for annual income and dummy variables for high earner and high education. See *Table A.1* for detailed variable definitions.

6 Summary and conclusion

Women are on average more absent from work for health reasons than men in most countries with sickness insurance systems. Surprisingly little research has been devoted to explaining the reasons behind this phenomenon, however. The few studies that do exist have mainly focused on factors such as differences in economic incentives, labor market characteristics, health and family responsibilities.

We take a different approach by seeking to investigate the potential role of behavioral differences in explaining the persistent gender gap in absenteeism. In particular, we first ask whether a relatively more risk-aversive and health-promoting behavior of women may serve as an explanation for the gender difference. In a subsequent second step, we investigate whether these preference differences can be attributed to utility-maximizing households seeking to maximize total household production in a setting where both lost market and home production are costly and home production is unevenly distributed within households. Moreover, as women have traditionally held dual roles as producers of both market and household goods, the potential increased benefit gained from investing in her health — through increased absence (recuperation) — would imply different absence patterns for men and women even when they enjoy the same level of health.

To investigate the potential role of gender-specific preferences in explaining the gender gap in absenteeism, we utilize administrative data from Sweden on sickness and disability benefits, mortality and in-hospital care. We argue that sickness absence is determined by both health and preferences while data on mortality and in-hospital care are more objective measures of individual health. As the gender difference in sickness absence may arise from unobserved factors correlated with health like e.g. productivity and economic incentives, estimates based on simple covariate adjustments will generally be confounded. For this reason, we sample men and women who have experienced a hospital admission, i.e. an observed health shock, and compare the relative change in sickness absence of men and women after the health change. By estimating the relative change in absenteeism we control for unobserved confounding factors provided that these unobserved factors are constant between groups over time. We show that if the magnitude of the health shock causing an admission to hospital is on average the same we can interpret the relative difference in post-admission sickness absence as a behavioral effect. According to the alternative hypothesis, in other words a more preventive behavior of women, the health shock requiring a visit would, however, be less severe for women than for men. The implication is that the estimated effect provides a lower bound of the behavior difference between women and men.

We find that womens' sickness absence increases after the admission to hospital compared to men. This result provides evidence that women display more preventative behavior when affected by deteriorating health. In addition, comparing the sickness absence of women with and without children we find that approximately one-third of this relative increase in sickness absence can be explained by women with children (i.e. individuals with a non-trivial amount of home production). Based on the analyses of both mortality and re-admission probabilities we find that the post-admission health of women is better than that of men and that household responsibilities are not related to worse post-admission health.

We perform an extensive number of sensitivity analyses. The results of the gender difference response of the admission are remarkably robust across all specifications. The results for the household analysis are somewhat more mixed, mainly due to more imprecisely estimated parameters. In particular, including disability beneficiaries in the work absence measure reduces the effect and the effect is only statistically significant when the outcome is the prevalence of sickness absence (rather than the annual number of days of absence).

In conclusion, the traditional morbidity-mortality paradox is commonly explained by the fact that women have innate precautionary and risk-aversive behaviors. The results obtained in this study lend support to the idea that differences in the health-related behavior of men and women are an important factor in explaining the gap in male-female absenteeism. Moreover, we find that a non-trivial share of these preference differences can be attributed to household investments in women's health.

We believe that the results obtained in this paper should be viewed as providing a more complex and multifaceted picture of the patterns of sickness absence and health across different demographic and socioeconomic groups. It would be a heroic task to attempt to isolate the effect of any particular health investment behavior to the gender difference in life expectancy. However, it is quite possible to argue in favour of this notion by noticing the negative correlation between the recent narrowing gender gap and the relative increase in female averse health behavior, such as obesity and smoking, observed in, for example, the U.K. (see LSAP (2012)). While the use of sickness absence as a means of investing in your health may not be on parity with the potential effects of losing weight or quitting smoking, it is probably still worth taking seriously in modern societies where stress-related diseases are emerging as a major health concern.

References

ÅKERLIND, I., ALEXANDERSSON, K., HENSING, G., LEIJON, M. and BJURULF, P. (1996). Sex differences in sickness absence in relation to parental status. *Scandinavian Journal of Public Health*, **24** (1), 27–35.

ANGELOV, N., JOHANSSON, P., LINDAHL, E. and LINDSTRÖM, E.-A. (2011). *Kvinnors och Mäns Sjukfrånvaro*. Rapport 2011:2, The Institute for Labor Market Policy Evaluation, Uppsala, Sweden.

ARRELÖV, B., EDLUND, C. and GOINE, H. (2006). Grindvakterna och sjukförsäkringen – samspel och motspel. In E. Palmer (ed.), *SKA Projektet: Sjukförsäkring, kulturer och attityder*, *10*, Försäkringskassan, Stockholm.

BERTRAND, M. (2010). New perspectives on gender. In O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics 4b*, *17*, Elsevier Ltd., pp. 1545–1592.

BOOTH, A. and VAN OURS, J. (2005). Hours of work and gender identity: Does part-time work make the family happier?, cEPR Discussion Paper No 5438, London.

BOYE, K. (2008). *Happy hour? Studies on well-being and time spent on paid and unpaid work*. Ph.D. thesis, Swedish Institute for Social Research, dissertation Series no. 74.

BRATBERG, E., DAHL, S.-Å. and RISA, A. E. (2002). The double burden – do combinations of career and family obligations increase sickness absence among women? *European Sociologial Review*, **18** (2), 233–249.

BRETT, G. (1969). Earlier diagnosis and survival in lung cancer study. *British Medical Journal*, **4**, 260–262.

BROSTRÖM, G., JOHANSSON, P. and PALME, M. (2004). Economic incentives and gender differences in work absence behavior. *Swedish Economic Policy Review*, **11**, 33–63.

BURDA, M., HAMERMESH, D. and WEIL, P. (2008). The distribution of total work in the eu and us. In T. Boeri, M. Burda and F. Kramarz (eds.), *Working Hours and Job Sharing in the EU and USA: Are Europeans Lazy? Or Americans Crazy?*, Oxford: Oxford University Press.

CROSON, R. and GNEEZY, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, **47** (2), 1–27.

ECKEL, C. and GROSSMAN, P. J. (2008). Sex and risk: Experimental evidence. In C. Plott and V. Smith (eds.), *Handbook of Experimental Economics Results vol 1*, Elsevier Ltd.

ENGLUND, L. (2001). *Förändringar i Distriktsläkares Sjukskrivningspraxis mellan åren 1996 och 2001 i ett Svenskt Landsting*. Working paper, Centrum för Klinisk Forskning Dalarna, Falun.

ENGSTRÖM, P. and JOHANSSON, P. (2012). The medical doctors as gatekeepers in the sickness insurance? *Applied Economics*, **44**, 3615–3625.

EVANS, O. and STEPTOE, A. (2002). The contribution of gender-role orientation, work factors and home stressors to psychological well-being and sicknessabsence in male- and female-dominated occupational groups. *Social Science and Medicine*, **54** (4), 481–92.

EVERTSSON, M. and NERMO, M. (2007). Changing resources and the division of housework: A longitudinal study of swedish couples. *European Sociological Review*, **23**, 455–470.

GRAMBSCH, P. M. and THERNEAU, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, **81** (3), 515–526.

GRAUNT, J. (1662). *Natural and Political Observations Mentioned in a following Index and made upon the Bills of Mortality*. London: Martin, Allestry and Dicas.

GROSSMAN, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, **80** (2), 223–255, the University of Chicago Press.

LAAKSONEN, M., MARTIKAINEN, P., RAHKONEN, O. and LAHELMA, E. (2008). Explanations for gender differences in sickness absence: evidence from middle-aged municipal employees from finland. *Occupational and Environmental Medicine*, **65** (5), 325–30.

LEE, C. (2010). Gender, health and health behaviors. In J. Chrisler and D. R. McCreary (eds.), *Handbook of Gender Research in Psychology*, *20*, Springer Science, pp. 471–493.

LEVIN, B. *et al.* (2008). Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: A joint guideline from the american cancer society, the us multi-society task force on colorectal cancer, and the american college of radiology. *Gastroenterology*, **134**, 1570–1595.

LSAP (2012). *Life Expectancy: Past and Future Variations by Gender in England and Wales*. Tech. rep., Longevity Science Advisory Panel, http://www.longevitypanel.co.uk/docs/life-expectancy-by-gender.pdf.

MASTEKAASA, A. (2000). Parenthood, gender and sickness absence. *Social Science and Medicine*, **50**, 1827–1842.

— and OLSEN, K. (1998). Gender, absenteeism and job characteristics: A fixed effects approach. *Work and Occupations*, **25**, 195–2228.

— and — (2000). Do women or men have less healthy jobs? an analysis of gender differences in sickness absence. *European Sociological Review*, **16** (3), 267–86.

NATHANSON, C. (1975). Illness and the feminine role: A theoretical review. *Social Science and Medicine*, **9** (2), 57–62.

PARINGER, L. (1983). Women and absenteeism: Health or economics? *American Economic Review*, **73** (2), 123–127, papers and Proceedings of the Ninety-Fifth Annual Meeting of the American Economic Association.

MCPHERSON, K. (2010). Should we screen for breast cancer? *British Medical Journal*, **341**, 233–235.

SCB (2009). *Tid för vardagsliv – Kvinnors och mäns tidsanvändning 1990/91 och 2000/01*. Report 99, Statistiska Centralbyrån, levnadsförhållanden.

SCHAPPERT, S. M. and NELSON, C. (1999). National ambulatory medical care survey: 1995-96 summary. *Vital and Health Statistics*, **142** (13), 1–122, http://www.cdc.gov/nchs/products/series.htm.

SELIN, H. (2009). The rise in female employment and the role of tax incentives. an empirical analysis of the swedish individual tax reform 1971, uppsala Center for Fiscal Studies Working Paper 2009:3.

SINDELAR, J. L. (1982). Differential use of medical care by sex. *Journal of Political Economy*, **90** (5), 1003–1019.

SOU (2004:43). Den könsuppdelade arbetsmarknaden. Statens offentliga utredningar (SOU). Stockholm 2004.

SSIA (2007:1). Försäkringskassan analyserar: Nej till sjukpenning. vad hände sen? Swedish Social Insurance Agency (SSIA). Stockholm 2007.

STRONEGGER, W.-J., FREIDL, W. and RASKY, V. (1997). Health behaviour and risk behaviour: Socioeconomic differences in an austrian rural county. *Social Science and Medicine*, **44** (3), 423–426.

TICHENOR, V. (1999). Status and income as gendered resources: The case of marital power. *Journal of Marriage and the Family*, **61**, 638–650.

UITENBROEK, D. G., KEREKOVSKA, A. and FESTCHIEVA, N. (1996). Health lifestyle behaviour and socio-demographic characteristics. a study of varna, glasgow and edinburgh. *Social Science and Medicine*, **43** (3), 367–377.

UK TRIAL OF EARLY DETECTION OF BREAST CANCER GROUP (1988). First results on mortality reduction in the uk trial of early detection of breast cancer. *The Lancet*, **332**, 411–416.

VERBRUGGE, L. M. (1982). Sex differentials in health. *Public Health Report*, **97** (5), 417–437.

Appendices

Appendix A. Tables and figures

Table A.1: Variable list

Variable	Description
Female	1 if female.
Child	1 if ever observed to have a child.
Age	The age in years of an individual.
Wage	Annual earnings in 2001 Basic Amounts.
High Earner	Annual earnings ¿ 7.5 Basic Amounts.
Income	Disposable income in 2001 Basic Amounts.
Virtual Income	The spouse's income in 2001 Basic Amounts.
Not Primary Earner	1 if earnings are less than the spouse's.
High Education	1 if tertiary level of education.
Sickdays	Annual number of insured days of sickness absence.
Shock	1 if a hospital admission.
Shock #2	1 if a second hospital admission.
Disability	1 if disability benefits.
Death	1 if deceased.

Table A.2: Means and standard deviations of the included variables by year

	Year												
Variable	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	Total
Female	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
Child	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82
Age	38.96	39.96	40.96	41.96	42.96	43.96	44.96	45.96	46.96	47.96	48.96	49.96	44.46
Wage	5.04	5.28	5.45	5.76	5.99	6.15	6.34	6.55	6.75	6.92	7.02	7.13	6.19
	(2.46)	(3.73)	(2.96)	(3.07)	(3.37)	(3.75)	(4.21)	(4.60)	(4.81)	(5.28)	(5.12)	(5.33)	(4.21)
High Earner	0.10	0.13	0.16	0.2	0.23	0.26	0.29	0.33	0.38	0.41	0.44	0.46	0.28
Income	8.26	8.69	9.06	9.61	10.04	10.40	10.9	11.53	12.21	12.73	13.14	12.33	10.72
	(4.17)	(5.38)	(4.89)	(5.29)	(5.79)	(6.25)	(13.34)	(25.47)	(39.78)	(48.93)	(58.03)	(8.67)	(26.37)
Virtual Income	3.22	3.41	3.61	3.85	4.05	4.24	4.56	4.98	5.46	5.81	6.12	5.19	4.53
	(3.48)	(3.91)	(3.90)	(4.21)	(4.56)	(4.80)	(12.54)	(24.99)	(39.46)	(48.63)	(57.80)	(6.21)	(25.97)
Not Primary Earner	0.31	0.31	0.32	0.32	0.32	0.32	0.33	0.33	0.33	0.33	0.33	0.32	0.32
High Education	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26
Sickdays	18.35	22.95	24.29	22.31	21.38	26.29	30.93	36.6	40.07	41.6	40.34	36.06	30.1
	(54.59)	(64.00)	(68.65)	(67.06)	(68.52)	(75.62)	(82.33)	(90.49)	(95.81)	(98.04)	(98.07)	(92.59)	(81.36)
Shock	0.19	0.19	0.18	0.17	0.16	0.15	0.14	0.14	0.13	0.13	0.13	0.13	0.15
Shock 2	0.00	0.04	0.08	0.13	0.17	0.21	0.25	0.29	0.32	0.36	0.39	0.43	0.22
Retired	0.02	0.02	0.03	0.04	0.04	0.05	0.06	0.07	0.09	0.10	0.12	0.14	0.07
Death	0.00	0.00	0.00	0.01	0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.03	0.02
Individuals	63,599	63,458	63,268	63,055	62,835	62,609	62,363	62,129	61,879	61,669	61,428	61,173	749,494

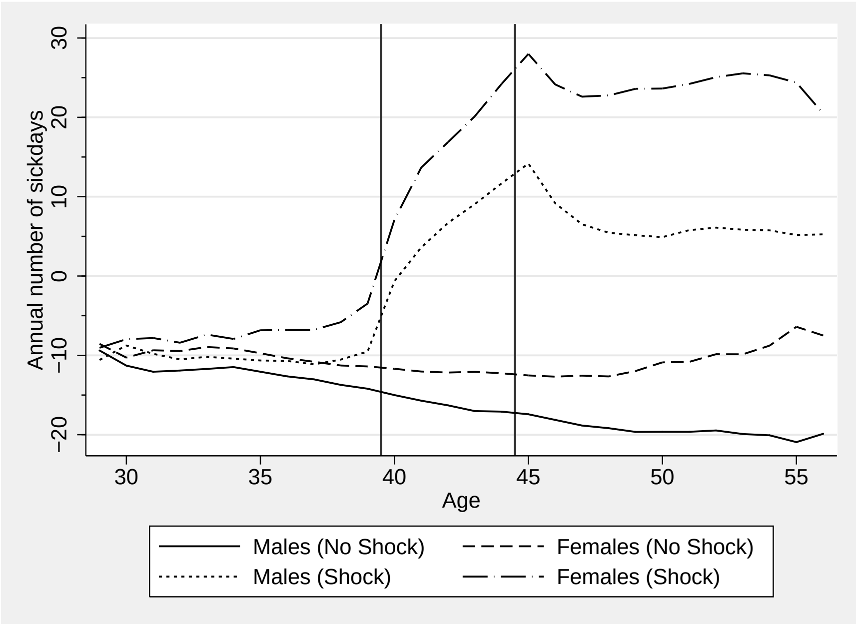
NOTE.—The table reports means and (standard deviations) of the analysis variables for the years 1993-2004. The summary statistics is based on the analysis sample defined in the data section. Incomes are measured in 2001 Swedish Basic Amounts (BA). One BA was approximately equal to 3,300i€n 2001. See *Table A.1* for detailed variable definitions.

Table A.3: Descriptive statistics for the analysis and the comparison samples

	Analysis sample				Comparison sample			
	Men		Women		Men		Women	
	Child	No Child	Child	No Child	Child	No Child	Child	No Child
Age	36.556	36.566	36.851	36.629	36.299	36.29	36.513	36.305
	(2.246)	(2.251)	(2.085)	(2.245)	(2.339)	(2.344)	(2.238)	(2.354)
Annual Earnings	7.05	6.567	4.62	5.676	7.435	6.855	4.83	5.931
	(3.993)	(2.526)	(1.96)	(2.142)	(4.147)	(2.645)	(2.127)	(2.06)
High Earner	0.31	0.243	0.062	0.132	0.353	0.278	0.077	0.156
	(0.463)	(0.429)	(0.24)	(0.338)	(0.478)	(0.448)	(0.266)	(0.363)
Disposable Income	5.267	4.68	4.338	4.172	5.412	4.845	4.326	4.241
	(3.82)	(1.72)	(3.858)	(2.014)	(3.03)	(1.637)	(1.64)	(1.386)
High education	0.209	0.171	0.316	0.332	0.241	0.195	0.348	0.355
	(0.406)	(0.376)	(0.465)	(0.471)	(0.428)	(0.396)	(0.476)	(0.479)
Not Primary Earner	0.072	0.011	0.573	0.09	0.067	0.014	0.61	0.083
	(0.259)	(0.105)	(0.495)	(0.286)	(0.25)	(0.116)	(0.488)	(0.276)
Virtual Income	2.674	0.34	4.963	0.905	2.858	0.375	5.318	0.85
	(2.548)	(1.391)	(7.907)	(2.571)	(3.094)	(1.675)	(4.655)	(2.524)
Sick Days	4.127	5.526	8.306	9.968	1.99	2.215	4.515	4.794
	(23.439)	(29.237)	(35.285)	(42.247)	(15.665)	(16.069)	(25.494)	(28.034)
Observations	58,015	12,105	49,670	8,113	75,373	14,204	52,586	8,573

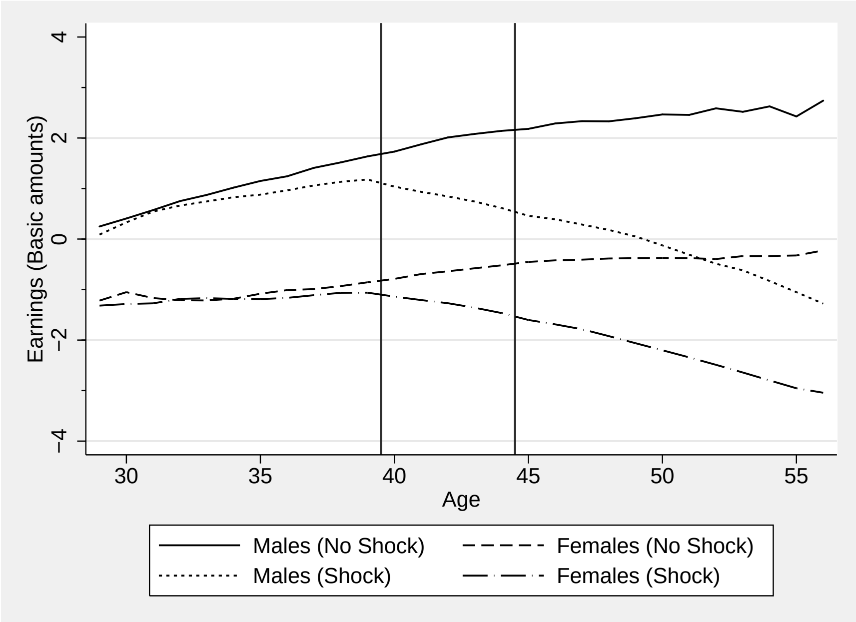
NOTE.—The table reports means and (standard deviations). The summary statistics is based on the share with age $<$ 40. The analysis sample consist of employed prime-aged individuals with the occurrance of a in-hospital care record sometime during ages 40-45. Each of these sampled individuals are matched with an employed prime-aged individual of the same age that did not have a hospital admission during the same age frame. For more information, see the data section in the paper. The summary statistics is disaggregated into gender and the presence of children. One BA (basic amount) in 2001 was approximately 3,300. € See *Table A.1* for detailed variable definitions.

Figure A.1: The average number of annual days of sickness absence for the analysis and comparison samples over age and by gender



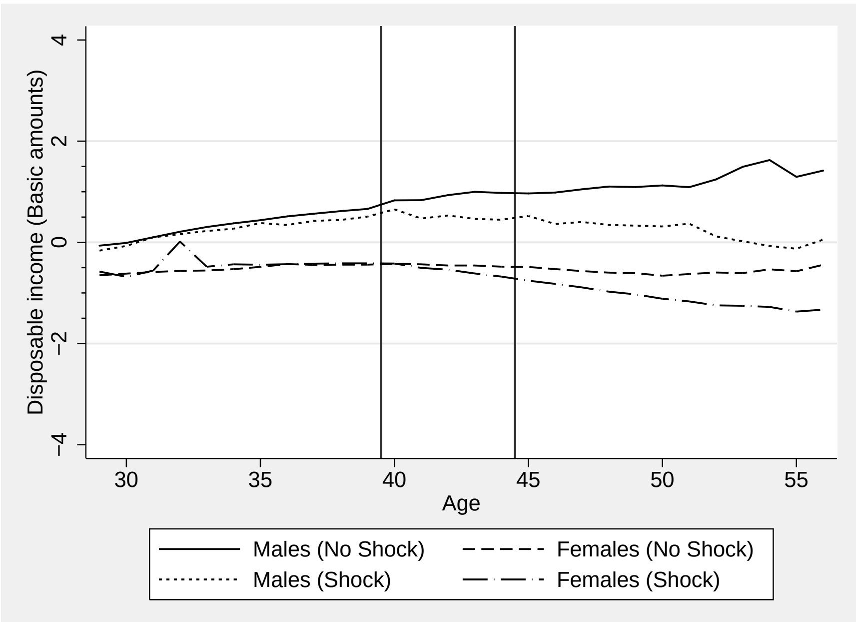
NOTE.—The figure is constructed by plotting the residuals from an ordinary least squares regression of year fixed effects on the outcome variable. Days of sickness are defined as the number of insured days an individual is observed to withdraw during one year. The analysis sample includes only individuals that had an observed hospital admission between the ages of 40 and 45 while the comparison sample did not have any registered hospitalization for the same age group. The "shock" ("No shock") lines pertain to the sample with (without) a hospital admission.

Figure A.2: The average annual earnings for the analysis and comparison samples over age and by gender



NOTE.—The figure is constructed by plotting the residuals from an ordinary least squares regression of year fixed effects on the outcome variable. Earnings are defined as gross cash earnings in the current year taken from the Swedish employment registry. The analysis sample includes only individuals that had an observed hospital admission between the ages of 40 and 45 while the comparison sample did not have any registered hospitalization for the same age group. The "shock" ("No shock") lines pertain to the sample with (without) a hospital admission.

Figure A.3: The average annual disposable income for the analysis and comparison samples over age and by gender



NOTE.—The figure is constructed by plotting the residuals from an ordinary least squares regression of year fixed effects on the outcome variable. Disposable income is defined as total individualized income in the current year net of tax, benefits and reductions. The analysis sample includes only individuals that had an observed hospital admission between the ages of 40 and 45 while the comparison sample did not have any registered hospitalization for the same age group. The "shock" ("No shock") lines pertain to the sample with (without) a hospital admission.

Table A.4: Share of females in the analysis sample by diagnostic category

Disease category	Number of females	Number of males	Total	Share of females
Accident	2,968	5,341	8,309	0.36
Blood	238	88	326	0.73
Cancer	5,085	985	6,070	0.84
Congential	150	127	277	0.54
Digestive	3,697	4,366	8,063	0.46
Ear	443	449	892	0.50
Endocrine	957	593	1,550	0.62
Eye	239	305	544	0.44
Factors	1,957	1,136	3,093	0.63
Genitourinary	5,084	1,231	6,315	0.81
Heart	1,710	2,899	4,609	0.37
Infection	651	941	1,592	0.41
Mental	1,402	1,783	3,185	0.44
Musculoskeletal	2,184	2,922	5,106	0.43
Nerve system	667	717	1,384	0.48
Perinatal	3	1	4	0.75
Respiratory	1,413	1,900	3,313	0.43
Skin	333	354	687	0.48
Symptoms	3,414	4,826	8,240	0.41
Total	32,595	30,964	63,559	0.51

NOTE.—The diagnosis categories are grouped according to the chapter division of the Swedish version of The International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10). Some of the categories does not pertain to specific diseases but are grouped for other reasons. These are; symptoms and signs of illnesses that cannot be classified differently (Symptoms), external causes such as injuries and poisoning (Accidents) and factors related to the contact with the medical establishment (Factors). Each cell value in columns (2)-(4) pertains to the number of sampled individuals with an registered hospital admission whose cause for the admission belonged to the specific category group at the time of the hospitalization. The female shares in the last column are the division of column (2) by column (4) for each row in the table.

Table A.5: Share of females in the analysis sample by industry sector

Industry sector	Number of females	Number of males	Total	Share of females
Agriculture	164	597	761	0.22
Construction	295	3,038	3,333	0.09
Education	3,225	1,486	4,711	0.68
Energy	125	486	611	0.20
Finance	782	594	1,37	0.57
Health	14,779	2,301	17,080	0.87
Services	406	316	722	0.56
Manufacturing	3,407	8,860	12,267	0.28
Mining	20	146	166	0.12
Other	123	288	411	0.30
Other Pers. Service	1,177	1,221	2,398	0.49
Public Administration	2,098	1,948	4,046	0.52
Real Estate and Renting	2,036	2,874	4,910	0.41
Retail and Wholesale	2,476	3,513	5,989	0.41
Transportation	1,482	3,296	4,778	0.31
Total	32,595	30,964	63,559	0.51

NOTE.—The table cover the two first digits of the industry sector code (SNI) aggregated into 15 categories covering the labour market. SNI is short for Swedish Standard Industrial Classification, which is closely based on the EU standard classification NACE Revision II. Each cell value in columns (2)-(4) pertains to the number of sampled individuals that are employed in november in the current year at an establishment classified within the specific industry sector in the first year they were sampled. As our sample only consist of employed individuals there are no missing sector codes for individuals in the sample. The female shares in the last column are the division of column (2) by column (4) for each row in the table.

Appendix B. Empirical modeling

This section derives the probability limits of the OLS estimator under different assumptions regarding the health shock. In particular, we show that under plausible assumptions the estimator of the gender difference in sickness absence will be conservative.

Gender differences

Assume labor supply and household production are fixed between two time periods, $t = (0,1)$, but that individual health is poorer in the second time period, i.e. $\eta_{i1j} - \eta_{i0j} = \Delta_{ij} < 0$; the change in the sickness absence equation (4) over the two time periods will be given by

$$s_{i1j} - s_{i0j} = \gamma_j \Delta_{ij} + \delta_f^{\#} c_i F_i \times \Delta_{ij} + \varepsilon_{i1j} - \varepsilon_{i0j},\ i = 1, ..., n_j, \tag{B.1}$$

where $n_j = n_{j1} + n_{j0}$ is the the number of men $(j = m)$ and women $(j = f)$ with $(c = 1)$ and without $(c = 0)$ children and with model parameters defined in Section 3. For women we have

$$s_{i1f} - s_{i0f} = \gamma_f \Delta_{if} + \delta_f^{\#} c_i \times \Delta_{if} + \varepsilon_{i0f} - \varepsilon_{i1f},\ i = 1, ..., n_f,$$

and for men

$$s_{i1m} - s_{i0m} = \gamma_m \Delta_{im} + \varepsilon_{i1m} - \varepsilon_{i0m},\ i = 1, ..., n_m.$$

From the i.i.d. of ε_{itj} we get

$$p\lim \frac{1}{n_f} \sum_{i=1}^{n_f} \left(s_{i1f} - s_{i0f} \right) = \gamma_f^{\#} \Delta_f \equiv \gamma_f^* > 0,$$

where $\gamma_f^{\#} = \gamma_f + \delta_f^{\#} p_{f1}$, $p_{f1} = n_{f1}/n_f$ is the fraction of women who have children and $p\lim \frac{1}{n_f} \sum_{i=1}^{n_f} \left(\eta_{i1f} - \eta_{i0f} \right) = (\overline{\eta}_{f1} - \overline{\eta}_{f0}) = \Delta_f < 0$ is the average health shock requiring an admission for women. For men we get

$$p\lim \frac{1}{n_m} \sum_{i=1}^{n_m} (s_{i1m} - s_{i0m}) = \gamma_m \Delta_m = \gamma_m^* > 0$$

where $p\lim \frac{1}{n_m}\sum_{i=1}^{n_m}\left(\eta_{i1f}-\eta_{i0f}\right)=(\overline{\eta}_{m1}-\overline{\eta}_{m0})=\Delta_m<0$ is the average health shock requiring an admission for men.

We can now specify

$$s_{it}=b_0+b_1F_i+\gamma_m^*H_{it}+\kappa^\circ F_iH_{it}+\varepsilon_{it}, \tag{B.2}$$

where

$$\kappa^\circ=\kappa\Delta_f+\gamma_m(\Delta_f-\Delta_m), \tag{B.3}$$

$$b_0=\phi_m\overline{C}_{sm}+\lambda_{m1}\overline{h}_m+\lambda_{m2}p_{m1}\overline{h}_f+\omega_{m2}p_{m1}\overline{h}_f^p+\omega_{m1}p_{m0}\overline{h}_m^p+\zeta_m p_m\overline{\mu}_m$$

and

$$b_1=\phi_f\overline{C}_{sf}+\lambda_{f1}\overline{h}_f+\lambda_{f2}p_{f1}\overline{h}_m+\omega_{f1}p_f\overline{h}_f^p+\zeta_f p_f\overline{\mu}_f-b_0$$

with parameters defined in Section 3, $p_j=n_j/n$, $p_{jc}=n_{jc}/n_j$ and $\overline{z}_j$ and $\overline{z}_{jc}$, $j=(m,f)$, $c=(0,1)$ are sample means and z being generic for (C_s,μ,h,h^p).

Assume that men an women have the same pre-admission health (i.e. $\overline{\eta}_{m0}=\overline{\eta}_{f0}$) and that the health requiring a hospital admission is on average the same for men and women.[18] The data generating process for the admission would then imply that $\tau_F=0$ in

$$H_i=\mathbf{1}(\eta_i<\tau+\tau_F F_i),$$

where $\mathbf{1}(\cdot)$ is the indicator function and τ is the average health threshold of being admitted at a hospital. This implies that $\Delta_m=\Delta_f=\Delta$. Under these two assumptions the ordinary least squares (OLS) estimator of equation (B.2) would be a consistent estimator of the average female to male difference in sickness absence from a negative health shock Δ,

[18]The hospital admission could be modelled as the combination of two different thresholds; one which governs the individual's decision to go to a hospital and one which governs the physician's decision to admit the visiting individual. We disregard from the latter threshold here, assuming that physicians are not able to perfectly screen the health status of individual patients.

that is

$$\kappa^* = \kappa\Delta.$$

Now, consider the situation where women invest more in their health than men. The health change requiring a hospital admission should thus be less severe for women on average, that is $\tau_F > 0$, or

$$\Delta_f - \Delta_m > 0.$$

Under this assumption the OLS estimator would be a biased estimator of κ^*. From the definition of κ° in equation (B.3) it follows that

$$p\lim \widehat{\kappa}^* - \kappa^* = \kappa\left(\Delta_f - \Delta\right) + \gamma_m\left(\Delta_f - \Delta_m\right),$$

where $\widehat{\kappa}^*$ is the OLS estimator. Under the alternative hypothesis, $\kappa < 0$, and $\gamma_m < 0$

$$p\lim \widehat{\kappa}^* - \kappa^* < 0.$$

The estimated gender effect of the health change on sickness absence will thus be conservative if women go to a hospital for less severe health problems than men.

In order to take any pre-admission differences in health between genders into account let

$$\eta_{i1j} = \eta_{i0j} + \omega_{i1j},$$

In this specification ω_{i1j} is the health shock from the original level η_{i0j}. Assume

$$\eta_{i0j} = X_i\beta + u_{ij},$$

where X_i is a vector of observed covariates, $\beta \neq 0$ and u_{ij} is i.i.d. This allows us to write the "structural" model (B.1) as

$$s_{i1j} - s_{i0j} = \gamma_j \omega_{i1j} + \delta_f^{\#} c_i F_i \times \omega_{i1j} + \varepsilon_{i1j} - \varepsilon_{i0j},\ i = 1, ..., n_j.$$

The implication is that the "difference in difference" specification (B.2) is given as

$$s_{it} = b_0 + b_1 F_i + \gamma_m^{\star} H_{it} + \kappa^{\star} F_i H_{it} + X_i \beta + \varepsilon_{it}\ i = 1, ..., n, \tag{B.4}$$

where

$$\kappa^{\star} = \kappa \overline{\omega}_{f1} + \gamma_m (\overline{\omega}_{f1} - \overline{\omega}_{m1}) \tag{B.5}$$

and $\overline{\omega}_{j1}$ is the average unobserved health shock between the two time periods for the males ($j = m$) and females ($j = f$), respectively. Under the assumption of equal average unobserved health shocks of men and women, the OLS estimator of equation (B.4) would converge to

$$\kappa_x^* = \kappa \overline{\omega}_1,$$

where hence $\overline{\omega}_1 = \overline{\omega}_{f1} = \overline{\omega}_{m1}$. If the unobserved shock requiring a hospital admission is on average less severe for women i.e. $(\overline{\omega}_{f1} - \overline{\omega}_{m1}) > 0$, the OLS estimator $\widehat{\kappa}_x^*$ is biased. It is evident from equation (B.5) that

$$p \lim \widehat{\kappa}_x^* - \kappa_x^* = \kappa(\overline{\omega}_{f1} - \overline{\omega}_1) + \gamma_m(\overline{\omega}_{f1} - \overline{\omega}_{m1}).$$

This will under the alternative hypothesis, $\kappa < 0$, again be biased downwards. If $\widehat{\kappa}_x^* > \widehat{\kappa}^*$, this suggests that women enjoy better pre-admission health than men.

Household investments

Let $w_{itjc} = c_i w_{itj} + (1 - c_i) w_{itj}$ for $w \in (s, \eta, \varepsilon)$. The difference in sickness absence over time for women with children is given by

$$\left(s_{i1f1} - s_{i0f1}\right) = \gamma_f (\eta_{i1f1} - \eta_{i0f1}) + \delta_f^{\#} c_i F_i \times (\eta_{i1f1} - \eta_{i0f1}) + \varepsilon_{i1f1} - \varepsilon_{i0f1}, i = 1, ..., n_{f1}, \tag{B.6}$$

Then, given i.i.d of ε_{itf1},

$$p\lim \frac{1}{n_{f1}} \sum_{i=1}^{n_{f1}} \left(s_{i1f1} - s_{i0f1}\right) = \left(\gamma_f + \delta_f^{\#}\right) \left(\overline{\eta}_{1f1} - \overline{\eta}_{0f1}\right) = \gamma_{f1}^{*},$$

where $p\lim \frac{1}{n_{f1}} \sum_{i=1}^{n_{f1}} (\eta_{i1f1} - \eta_{i0f1}) = (\overline{\eta}_{1f1} - \overline{\eta}_{0f1}) = \Delta_{f1}$. Similarly, for women without children the difference is given by

$$\left(s_{i1f0} - s_{i0f0}\right) = \gamma_f (\eta_{i1f0} - \eta_{i0f0}) + \varepsilon_{i1f0} - \varepsilon_{i0f0}, i = 1, ..., n_{f0},$$

where n_{f1} is the number of females without children. Now

$$p\lim \frac{1}{n_{f0}} \sum_{i=1}^{n_{f0}} \left(s_{i1f0} - s_{i0f0}\right) = \gamma_f (\overline{\eta}_{1f0} - \overline{\eta}_{0f0}) = \gamma_{f0}^{*},$$

where $p\lim \frac{1}{n_{f0}} \sum_{i=1}^{n_{f0}} (\eta_{i1f0} - \eta_{i0f0}) = (\overline{\eta}_{1f0} - \overline{\eta}_{0f0}) = \Delta_{f0}$.

We can now use the same type of "difference in difference" specification as for the gender comparison in the analysis in the OLS estimation; that is

$$s_{it} = b_0^* + b_1^* c_i + \gamma_{f0}^* H_{it} + \delta^{\circ} c_i H_{it} + \varepsilon_{it},\ i = 1, ..., n_f, \tag{B.7}$$

where

$$\begin{aligned} \delta^\circ &= (\delta_f^\# + \gamma_f)\Delta_{f1} - \gamma_f \Delta_{f0}, \qquad \text{(B.8)} \\ b_0^* &= \phi_f \overline{C}_{sf} + \lambda_{f1}\overline{h}_f + \omega_{f1}\overline{h}_f^p + \zeta_m \overline{\mu}_m \end{aligned}$$

and

$$\begin{aligned} b_1^* = {} & \phi_f(\overline{C}_{sf1} - \overline{C}_{sf}) + \lambda_{f1}(\overline{h}_{f1} - \overline{h}_f) \\ & + \omega_{f1}(\overline{h}_{f1}^p - \overline{h}_f^p) + \zeta_m(\overline{\mu}_{m1} - \overline{\mu}_m) + \lambda_{m1}\overline{h}_m \end{aligned}$$

where $\overline{C}_{sfc}, \overline{h}_{fc}^p$, and $\overline{h}_{fc}$ are mean values for women with children ($c = 1$) and without children ($c = 0$).

If the health shock that leads to a hospitalization is the same for both groups, i.e. $\Delta_f = \Delta_{f0} = \Delta_{f1}$,

$$\delta^\circ = \delta^* = \delta_{f1}^\# \Delta_f.$$

If the the pre-admission health of women with and without children is the same, but the health shock for women with children rendering a hospital admission is lower than for women without children, that is

$$\Delta_{f1} - \Delta_{f0} > 0,$$

the OLS estimator of δ^* will be biased downwards. From equation (B.8) we get

$$p\lim \widehat{\delta}^* - \delta^* = (\delta_f^\# + \gamma_f)\left(\Delta_{f1} - \Delta_f\right) - \gamma_f \Delta_{f0}$$

and under the assumption, $\Delta_{f1} - \Delta_{f0} > 0$, $p\lim \widehat{\delta}^* - \delta^* < 0$.

Assume $\eta_{i1jc} = \eta_{ijc0} + \omega_{i1jc}$ and that initial health follows $\eta_{i0jc} = X_i\beta + u_i,$. This allows us to write the structural model (B.6) as

$$\left(s_{i1fc} - s_{i0fc}\right) = \gamma_f \omega_{i1jc} + (\delta_f^{\#} + \gamma_f) c_i F_i \times \omega_{i1jc} + \varepsilon_{i1fc} - \varepsilon_{i0fc}, i = 1, ..., n_{fc},$$

The implication is that the "difference in difference" specification (B.7) is given as

$$s_{it} = b_0^* + b_1^* c_i + \gamma_{f0}^* H_{it} + \delta_x^* c_i H_{it} + X_i \beta + \varepsilon_{it}.$$

Now

$$\delta_x^* = (\delta_f^{\#} + \gamma_{f0}) \overline{\omega}_{1f1} - \gamma_{f0} \overline{\omega}_{1f0}, \tag{B.9}$$

where $\overline{\omega}_{1fc} < 0$ is the average unobserved health shock in period one for women with ($c = 1$) and without ($c = 0$) children. Under the assumption of equal average unobserved conditional health shocks we have

$$\delta_x^* = \delta_f^{\#} \overline{\omega}_{f1},$$

since $\overline{\omega}_{f1} = \overline{\omega}_{1f1} = \overline{\omega}_{1f0}$. If the unobserved shock requiring a hospital admission is on average less severe for women with children, i.e. $(\overline{\omega}_{1f1} - \overline{\omega}_{1f0}) > 0$, the OLS estimator $\widehat{\delta}_x^*$ is biased. It is evident from equation (B.9) that

$$p \lim \widehat{\delta}_x^* - \delta_x^* = (\delta_f^{\#} + \gamma_{f0})(\overline{\omega}_{1f1} - \overline{\omega}_{1f}) - \gamma_{f0} \overline{\omega}_{1f0}.$$

which, under the alternative hypothesis that $\delta_{f1}^{\#} < 0$, will again be biased downwards. If $\widehat{\delta}_x^* > \widehat{\delta}^*$, this supports the idea that women with children have better pre-admission health than those without.

Assume the household differences in the magnitudes of the health shocks between men and women are proportional, so that

$$\Delta_{f1} - \Delta_{f0} = \pi \left[\Delta_{m1} - \Delta_{m0} \right],$$

where $\Delta_{m1} = (\overline{\eta}_{m11} - \overline{\eta}_{m10})$, $\Delta_{m0} = (\overline{\eta}_{m01} - \overline{\eta}_{m00})$, and where $\pi = \gamma_m / \gamma_f < 1$ under the alternative hypothesis, $|\gamma_f| > |\gamma_m|$. Under this assumption men with and without children can be used to control for the difference in the threshold for making a hospital admission among women with and without children. The implication of the assumption is that the health threshold difference for hospital admittance should be greater for men with and without children than the corresponding difference for women with and without children. Based on the information contained in *Figure 5.4* we believe this assumption to be plausible. The OLS estimator of δ^* using a triple difference model is subsequently computed from

$$\begin{aligned} s_{it} &= b_0 + b_1 F_i + b_2 c_i + b_3 F_i c_i + \gamma^*_{m0} H_{it} + \gamma^*_{m1} H_{it} c_i \\ &\quad + \gamma^*_{f0} F_i H_{it} + \delta^* F_i H_{it} c_i + u_{it} \end{aligned}$$

where γ^*_{m1} and γ^*_{m0} is the response of men with and without children.