

Schön, Matthias

Conference Paper

Unemployment, Sick Leave and Health

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Health Economics 2, No. B19-V3

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Schön, Matthias (2015) : Unemployment, Sick Leave and Health, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Health Economics 2, No. B19-V3, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at:

<https://hdl.handle.net/10419/113013>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Unemployment, Sick Leave and Health*

February 27, 2015

Abstract

This paper studies the relationship between sick leave, income and unemployment. In particular, it investigates this relationship under the generous German sick leave regulation of 100% wage replacement, i.e., in an environment where workers do not bear any direct costs from missing work due to sickness. Using information from the German Socioeconomic Panel (GSOEP) I identify three stylized facts of sick leave in Germany. First, sick days show a strong pro-cyclical pattern. Second, average use of sick days is hump-shaped over income quintiles. Third, the number of sick days is a strong predictor of becoming unemployed. Using this micro-evidence I develop a structural model that rationalizes these facts. I argue that in absence of direct costs of sick leave the fear of future unemployment is the main driving force restraining sick leave. I then use the model to do counterfactual policy analysis.

JEL Classification: D31, E24, J22, J32, H51, I14, I18.

Keywords: Health Economics, Unemployment, Sick Leave, Inequality

*Part of this project has been realized while visiting the economics department of the University of Pennsylvania. I am indebted to Dirk Krueger for his invitation. I thank seminar participants at the iHEA Congress, the University of Cologne, the University of Frankfurt, Hanming Fang, David Jaeger, and Alexander Ludwig for helpful comments.

1 Introduction

Losses of labor income due to sick leave are one major economic risk for the working age population associated with health. Direct costs arise from a reduction of time worked which would be paid otherwise. The extent to which these costs are insured varies greatly across industrialized countries, cf. Scheil-Adlung and Sandner (2010). An extreme case is the US with no statutory paid sick leave - a situation US president Barack Obama recently urged to overcome in his State of the Union 2015.¹ On the other extreme, Germany has one of the most generous sick payment systems. In case of work absence due to sickness, every (full-time, part-time or temporary) employee is eligible for six weeks of 100% of wage replacement.² This generosity comes with a price. The expenditures of paid sick leave which are borne by employers amount to almost 40 billion € in 2013 or more than 1.5% of GDP according to the German Federal Ministry of Labour and Social Affairs, cf. Bundesministerium für Arbeit und Soziales (2014).³ Indirect costs of sick leave stem from reductions in future expected earnings. Layoff or promotion decisions of employers depend on the past sick days of workers, cf. Markussen (2012) and further evidence below.

In this paper, I primarily focus on the indirect costs of sick leave.⁴ To this end, the first objective is to identify and document patterns of sick leave utilization over business cycles, income quintiles and age groups in Germany. The second objective is to rationalize these empirical findings within a theoretical framework and highlight the main mechanism – the decision to trade off utility-enhancing health against expected future earnings due to increased risk of job loss. The third objective is to analyze the distributional effects of indirect costs sick leave and evaluate counterfactual policies within a structural model calibrated to the German labor market.

To the first objective of this paper, I employ data from the German Socioeconomic Panel (GSOEP), a nationally representative longitudinal data set. With respect to aggregated data I find three remarkable patterns of sick leave utilization in Germany. First, average claims of sick leave exhibit a strong pro-cyclical pattern with a correlation of -0.7155, i.e., workers are on average less absent in times of high unemployment. Second, average sick days in Germany display a marked hump-shaped pattern over income quintiles. Workers in the medium income quintile have on average 10% more sick days than workers in the bottom income quintile. This is noteworthy because average health is monotonically increasing in income. Assuming sick days are only driven by health would therefore lead to a decrease in the number of sick days between bottom and medium

¹Paid sick leave may be provided by the employer on a contractual basis. According to the State of the Union the number of workers without any sick leave scheme however amounts to 43 billion.

²For more information on Germany regulation on sick leave see Appendix 7.1.

³This number does not include the contribution to the social insurance system which additionally amount to 6.9 billion € .

⁴This is possible as the German system of full wage replacement ruling out direct costs of sick leave.

income quintile. This is not observed. Third, the variance of sick days differs greatly between income quintiles. Employees in the bottom quintile have the highest probability to not miss any day a year. They also have the highest probability to miss more than two weeks. Top income employees miss small number of sick days but on a high frequency. Exploiting the panel structure of the GSOEP I additionally show that sick days are one key predictor of future unemployment. Taking five additional sick days increases the risk of becoming unemployed by 10%.

As to the second objective, the mechanism that rationalizes these sick leave patterns is that workers who become sick face the decision to either stay at home and recover or going sick to work. Staying at home restores utility-enhancing health but at the same time increases the risk of job loss. Going to work sick preserves expected future earnings but a perpetual neglect of recuperation diminish long-run health prospects.⁵ In times of a high unemployment rate, workers face both higher overall firing rates as well as lower reemployment probabilities. Resulting higher marginal costs of unemployment shift the trade-off towards presenteeism and drive the cyclical pattern. Workers facing financial constraints, i.e., low skilled workers, are less able to smooth consumption over periods of unemployment and have an overall higher risk to become unemployed. Therefore they are particular compelled to go to work when sick. In the end, optimal sick day utilization differs between income groups. Rich take constantly small number of sick days to conserve health. Low skilled worker reduce sick days to keep their job. Resulting low health increases the however small probability to be hit by a severe shock and resulting high number of sick days.

To quantify the distributional effects of sick leave, the third objective, I develop a heterogeneous agent model with endogenous health and incomplete credit and insurance markets. Additionally, I implement central characteristics of the German health care and worker protection system into my model. Individuals are entitled to continued wage payments for up to six weeks per sickness episode.⁶ The government imposes the progressive German income tax schedule on agents. The collected revenues are used to finance (i) expenditures due to sick leave payments (ii) unemployment benefits, (iii) a retirement system.

To empirically implement my quantitative analysis I first estimate and calibrate the model using GSOEP data to match key statistics on sick leave, health status and unemployment. In my calibration strategy I set some of the parameter values outside of the model (e.g., the interest rate, preference parameters, policy parameters, etc.). The income process, the layoff probability and incidence of sick days are estimated directly from the data. For remaining parameters (e.g., health transition probabilities, etc.) I use

⁵There exists a growing medical literature on the negative effect of neglect of recuperation time, cf. Kivimäki et al. (2005), Bergström et al. (2009).

⁶Individuals are also offered universal health insurance that covers medical expenditures. For this purpose I omit medical expenditures in my model.

my model to choose their values.

The estimated model is able to successfully explain the targeted features of the data in the estimation (e.g., cyclicalities of claims of sick leave, decrease of health over the life cycle, and the different utilization of sick days by low and high skilled). It is also capable to explain other (non-targeted) dimensions such as the different pattern in intensive and extensive margin of sick days as well as the income gradient in health.⁷

I then use the parameterized version of the model as a laboratory to evaluate the consequences of different policy options. To this purpose, I contrast the benchmark economy including paid sick leave with full wage replacement (Germany) with no mandatory paid sick leave (US), a wage independent amount of paid sick leave (UK) and a system with an unpaid grace period.

Related Literature The empirical part of this paper is related to two strands of the literature. First, this paper confirms for Germany the well documented negative relationship between the economic situation and the average claim of sick days per worker, starting with Leigh (1985). This literature shows that there are two potential mechanisms regarding the negative relationship between the unemployment rate and average sick days. One is an incentive effect, i.e., unemployment affects the propensity to take sick days. If an employee's higher sick rate increases the risk of job loss, a higher unemployment rate reduces the propensity to take sick days. A second alternative mechanism is related to the absence behavior of marginal workers entering or leaving the working population in various states of the business cycle. When employers can choose whom to layoff, the most absence-prone workers are more likely to be laid off in an economic downturn. All agree that this pattern is mainly driven by the reduction of sick time of workers due to fear off job loss in recessions. Arai and Thoursie (2005) and Askildsen et al. (2005) show that the incentive effect is the dominant force.

Second, I provide additional evidence for this incentive effect on a micro basis, i.e., I estimate the impact of sick days on future layoff risk. Other work in this area is done by Hesselius (2007). They find that sick days are a strong predictor off job loss. Andersen (2010) finds that many sick days not just affect employment risk but also decrease post sick leave earnings.

The structural model I present in this paper is part of a broad and growing macro health literature that incorporates endogenous health into dynamic models. Important related contributions include Grossman (1972), Ehrlich and Chuma (1990), Hall and Jones (2007), Ales et al. (2012), Halliday et al. (2012), Ozkan (2011), Cole et al. (2014).

Only a small literature distinguishes in such dynamic models between long run health and the onset of acute illnesses. Gilleskie (1998) predicts the change in physician services

⁷For example, in excellent health in 1984 have 74 percent more wealth than respondents in fair or poor health do, cf. Smith (1999).

use and illness-related absenteeism that arise with improvements in access to health care through more complete health insurance and sick leave coverage in the US.⁸ The paper however, only focus on the direct costs of work absence and does not take the risk of unemployment into account. It also falls short to provide a link to the endogenous health literature.

The rest of the paper is organized as follows: In Section 2, I discuss the main data source, the methodology and the empirical findings. Then I introduce a full structural model in Section 3. In Section 4.1, I discuss the estimation of the model and the model's fit to the data. Then I perform counter-factual policy experiments using the model in Section 5. Finally, I conclude in Section 6.

2 Empirical Facts

The purpose of this section is to motivate the key modeling assumption of the structural model in section 3. After discussing the data source and the methodology in section 2.1 I present in section 2.2 findings on aggregated data that proof that sick days are an endogenous choice of workers. Then, in section 2.3 I present results based on a panel analysis that underline the importance of the sick day choice for future income of workers.

2.1 Data and Methodology

2.1.1 Description of the Survey

My empirical analysis is based on the German Socio-Economic Panel (GSOEP), a nationally representative longitudinal data set. Starting in 1984 and conducted annually and it comprises 30 waves of data. It oversamples foreigners, immigrants, and East Germans to allow for more precise estimates for population subgroups that may be of particular policy interest.⁹ GSOEP provides detailed information about demographic (sex, age, ...), socioeconomic (educational level, marital status, ...) and economic characteristics. The respondents report their monthly income in the current and the previous year.¹⁰ The employment history contains the current employment status (full time, part time...), point in time of layoff in previous year, length off unemployment spell and information about the time worked for the same firm. Information about health is asked since 1990. The GSOEP contains information about self-reported health, number of doctor visits and

⁸Health-enhancing leisure in the health economics literature goes back to e.g., Grossman (1972), Ruhm (2000), while clinical, experimental, and empirical evidence in support of this idea can be found in the bio-medical science, public health, psycho-biology, and bio-sociology, and empirical health literature.

⁹I include all sub-samples of GSOEP with the appropriate cross-sectional weights.

¹⁰Both income variables are deflated with the consumer price index contained in the GSOEP using 2005 as base year.

hospital stays.¹¹ Further detailed information about the characteristics of the GSOEP is provided in Wagner et al. (2007).

Key variable for the purpose of this paper is the number of working days missed due to sickness. The GSOEP asks the respondents to state if they missed any day due to sickness in the previous year and if so how many. Puhani and Sonderhof (2010) show that, though self-reported, the GSOEP adequately depicts the true number of days of absence from work. The GSOEP also contains information about the number of spells that last longer than six weeks. However, the survey does not record information about occasion respondents going sick to work.

The only information that I used and is not contained in the GSOEP is the unemployment rate of Germany. I use official data from the federal employment agency, cf. Bundesagentur für Arbeit (2014). To investigate the cyclical behavior I additionally construct a dummy variable indicating periods with a high rate of unemployment, "Recession".¹²

2.1.2 Determination of the Sample

For the following empirical results not all of the observations of GSOEP are used. As I am interested on the sick leave use of workers I focus only on the working age population. I drop all observations of respondents younger than 18 and older than 65 (official German retirement age). I restrict the sample to respondents that either report to work in the current or in the previous year or report to be unemployed. This excludes people that are doing their military or social service as well as people that report not to be employed and not looking for work. I also exclude part time worker and respondents that report a monthly income of less than 400 €. ¹³ The probability and the intensity of annual sick leave are biased when respondents only work a fraction of the year.

As time period I use the waves 1994 to 2012, corresponding to information about sick days from 1993 to 2011. Waves 1984, 1990 and 1993 do not contain information about sick days. Wave 1991 and 1992 captures the unique economic situation of German reunification in 1990 and the liberalization of a state-owned socialist economy. I dropped both periods as income distribution and employment situation changed dramatically.¹⁴ Waves 1985 - 1989 could potentially be used in the analysis of pro-cyclicality. I drop them for various reasons. First, these waves do not contain information about health and

¹¹It also includes a SF-12 indicator of physical health. This measure combines several self-reported indicators, see Nuebling et al. (2007) for further information. Unfortunately, this measure is only available every second year since 2002 and is only of limit use for the panel analysis.

¹²For the results in Table 7 set an unemployment rate of 11% as cut off level. This gives me equally many periods marked as recessions as not. The following results are robust for other cut off levels.

¹³All monetary variables are deflated with the base year 2005.

¹⁴Following waves are also affected but the effect is strongly mitigated over time. Especially the classification in income quintiles is disturbed as the income scale was lower in East-Germany. **Still to do:** Do the analysis for West Germany only.

cannot be used in cross section or panel regressions. Second, the unemployment rate only varied in these year between 7.9% and 8.1%. Hence, not much variation can be used for figuring out cyclical of sick leave. And third, I want to use an uninterrupted sample period for the time series analysis.

For the benchmark results I exclude civil servants and self-employed from the sample. Self-employed workers do not receive paid sick leave as it is provided by the employer. Civil servants do get paid sick leave but are not eligible for layoffs. Hence, they are not or less affected by the indirect effect of fear of job loss.¹⁵

Sick days have a highly skewed distribution with many observations on the 0-boundary and few observations at the highest value of 365. 95% percent of the observations report less than 40 sick days a year and only one percent report more than 120 days. Hence, many results, e.g., the average number of sick days, are prone to be driven by only a few observations. To control for outliers I exclude in the benchmark results all observations that have one or more spells of sick days that last longer than six weeks. Of the remaining sample I cut off the highest one percent, i.e., workers with more than 40 sick days a year.

After sample selection, the sample used for benchmark results consists of 145,156 observations.¹⁶ It includes 20 waves and each wave has at least 6,160 observations.¹⁷

2.1.3 Empirical Approach

In Section 2.2 I run cross section regressions of sick days on various regressors. The regression equation for OLS and Logit are

$$S_i = \alpha + \beta \log(W_i) + \beta H_i + \beta \bar{U}_i + X_i\theta + \varepsilon_i \quad (1)$$

$$\text{Logit} [S_i^{\text{ext}} = 1] = \Phi\{\alpha + \beta \log(W_i) + \beta H_i + \beta \bar{U}_i + X_i\theta + \varepsilon_i\} \quad (2)$$

where S_i is a countable variable denoted for sick days whereas S_i^{ext} is a dummy variable for either miss any day ($S_i^{\text{ext}} = 1$) a year or not ($S_i^{\text{ext}} = 0$). W_i is the monthly income of the respondent in the previous year.¹⁸ H_i is self-reported health, $\bar{u}_i$ the unemployment rate and $X_{i,t}$ is a set of control variables, e.g., sex, age, years of education, year dummies and ε_i is the random error term.

In Section 2.3 I employ a logistic panel regression. I estimate the effect that (accumu-

¹⁵Results of both groups are an additional argument for the proposed mechanism. The cyclical pattern is either not existent for the self-employed or much weaker for the civil servants, a result also found by Pfeifer (2013). Also the income gradient in sick days does not exist for both groups. Unfortunately both groups vary from the rest of the sample in various respects (e.g., income, age, education). Therefore they cannot be used as an adequate control group.

¹⁶Further details of the sample selection can be found in Appendix 7.2.

¹⁷The size of the waves increases over time. There were refreshments of the GSOEP in 1998, 2000, 2002 and 2006.

¹⁸I used other measures of income that are also included in the GSOEP. The qualitative results do not change.

lated) sick days in the previous period have on the current probability of unemployment. I restrict the sample in this section to people that were employed at least six month in the previous year. The panel structure of GSOEP additionally allows me to use a fixed effects model. The fixed effect will incorporate all unobserved characteristics of the agent. The regression equations are

$$\text{Logit} [\text{Unemployment}_{i,t} = U] = \Phi\{\alpha + \beta S_{i,t-1} + C_{i,t-1}\theta + \epsilon_{i,t}\} \quad (3)$$

$$\text{Logit} [\text{Unemployment}_{i,t} = U] = \Phi\{\alpha + \beta S_{i,t-1} + C_{i,t-1}\theta + a_i + \epsilon_{i,t}\} \quad (4)$$

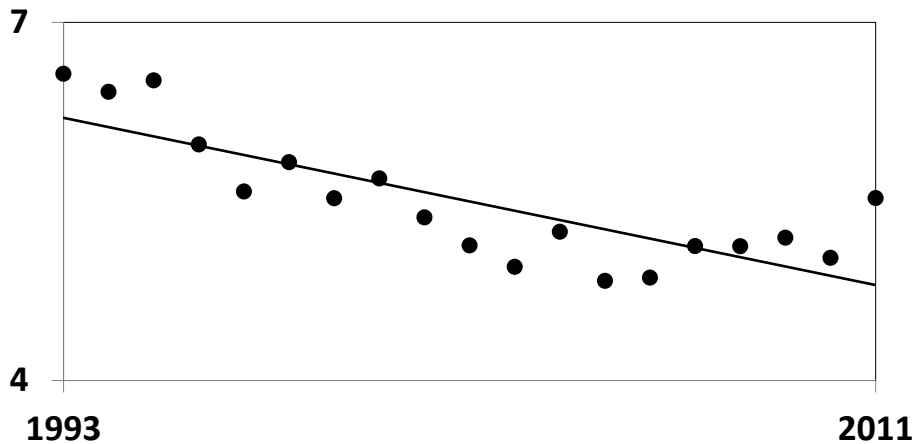
where $C_{i,t}$ is a set of control variable that do vary over time. It contains lagged health, age and lagged log income. For sick days, $S_{i,t-1}$, I use two different definitions. First, sick days reported only in the previous year. Second, accumulated sick days, i.e., the average sick days of the respondent over the last three years. The a_i s represent the individual specific and time-invariant fixed effect component and $\epsilon_{i,t}$ is the random error term.

2.2 Facts on Aggregated Data

2.2.1 Time series

Figure 1 shows for the benchmark sample the average annual number of sick days of workers in the observed time period and a fitted linear trend. Average sick days varies between 4.8 and 6.6 days. An obvious first finding about the number of sick days in Germany is the long term decline. In the last 19 years average claims of sick days have declined by more than 1.4 days or 20% relative to 1993.¹⁹

Figure 1: Average Number of Sick Days per Worker 1993 - 2010

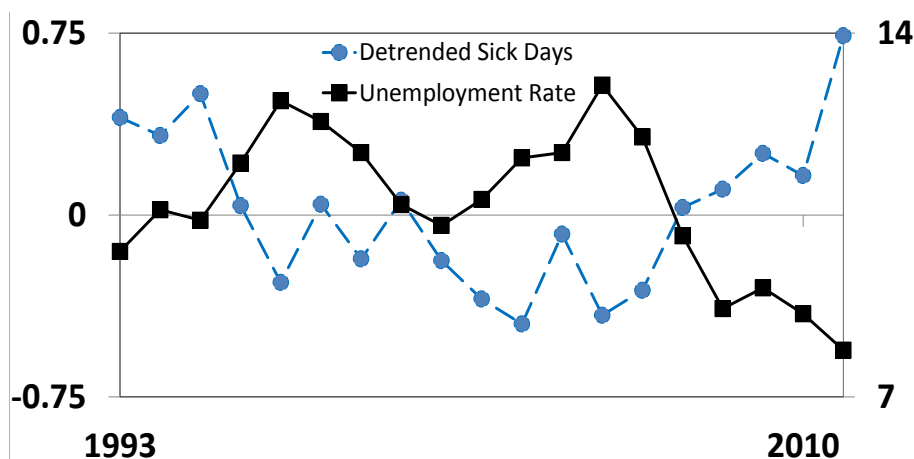


Notes: Dots: Average annual claims of sick days for benchmark sample. Solid line: Fitted linear trend.

¹⁹A potential explanation might be pressure from abroad on the German labor market. Another explanation might be that the decline is due to technological progress in medical treatment.

A second empirical characteristic is the strong pro-cyclical pattern of average claims of sick leave in Germany once the time series is de-trended.²⁰ Figure 2 depicts the absolute deviation of average number of sick days from the linear trend (dashed line) and the unemployment rate for Germany (solid line). Average number of sick days is high when the German unemployment rate is low. The correlation between the de-trended time series of sick days and the German unemployment rate is for the benchmark sample -0.7155 .²¹

Figure 2: Deviation of Sick Days from Linear Trend and Unemployment Rate



Notes: Dashed line (left axis) is absolute deviation of average number of sick days from the linear trend for benchmark sample; solid line (right axis) is the German unemployment rate.

To control for the composition effect, i.e. the absence behavior of the marginal worker, I construct a sub-sample consisting of workers that never report to be unemployed and have been observed for at least five consecutive years. This sub-sample shows on the one hand a lower number of average sick days compared to the benchmark sample. On the other hand the cyclical pattern of always-employed sample is still distinctive negative with a correlation of -0.6636 .²² The remaining correlation supports the incentive effect, i.e., in times of low reemployment workers reduce their number of sick days to avoid unemployment. The incentive effect implicitly assumes that absence from work is not mechanically tied to the incidence of sickness. Workers are free to decide whether to go to work sick or stay at home and recover. This is a key assumption of the structural model in section 3.

²⁰Pro-cyclicality of sick days is also documented for other countries, see Leigh (1985), Askildsen et al. (2005).

²¹In Appendix 7.3 I provide additional robustness checks using other measure of central tendencies, e.g., the number of sick days for the median worker.

²²Other composition effects would occur if specific occupation groups or sectors that exhibit high sick days, e.g. the construction sector, are hit stronger by business cycles than others. In Appendix 7.3 I show that the general pattern of pro-cyclical behavior holds for all occupation and sectors.

2.2.2 Cross Section

The economic trade-off between taking sick days and recover on the one side and increased layoff probability on the other side is also shown in another pattern of sick days. A cross sectional analyses of sick days exhibits remarkable differences of average sick leave use between income groups.²³ Figure 3 plots the average number of sick days (solid line) for each income quintile. It additionally shows the average self-reported health (dashed line) for each quintile. The used sample is restricted to workers between age 40 to 50. Controlling for age is important as age is highly correlated with sick days, health and income. Pooling all observations would lead to a bias in the results as poor people are more likely to be young and therefore healthy and using less sick days.²⁴ Workers in the top income quintile have the lowest number of average sick days. Workers in the medium income quintile claim on average more sick days than workers in the bottom income quintile.²⁵ On the contrary, the health profile over income quintiles is monotonically increasing. The poorest workers have the lowest average health where the top income quintile shows the highest average health.²⁶

Health and sick days are naturally related, i.e., worker in bad health are more likely to be sick and stay at home. Differences in health could potentially explain the small use of sick leave in the top income quintile compared to the rest of the workforce. Rich people are on average less sick and do not need to stay at home to recover. However, the same rationale is puzzling on the other side of the income distribution. The ones that are most unhealthy use less sick days to recover than medium income workers that have on average a better health. This is further evidence that sick days are not mechanically tied to health and absence behavior of workers has a second determinant.

This graphical inspection is confirmed by estimating Equation (1) using number of sick days as dependent variable. Both income coefficients, for log income and for log income squared, are highly significant and suggest a hump shaped relationship of income and sick days. Health has the assumed protective effect against sick days. Other coefficients in Table 1 confirm former results. There is a long run negative trend in sick days of -0.0867 days per year. More importantly relating to the cyclicalities of sick days, the coefficient of the unemployment rate is significantly negative. This means during periods of high unemployment average number of sick days are reduced.

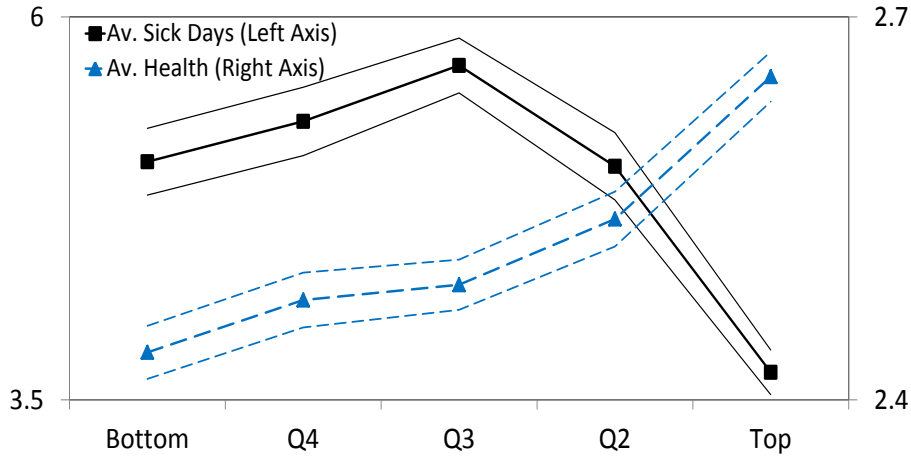
²³ Respondents are classified into income quintiles based on their monthly income in the previous year. If they report to be unemployed in the last year they are not considered for this section as they also have no sick days.

²⁴ All results hold also for other age bins, see Figure 12 in Appendix 7.3. See Appendix 7.3 for a version of these figures where gender effects are controlled for. The qualitative results are not changing.

²⁵ This pattern also exists for the median number of sick days in each income quintile and other sick days cut off levels.

²⁶ A simple probit regression of income on health (good or bad) controlling for age and sex confirms this pattern and yields a highly significant positive effect for log income. This income gradient in health is well established in the literature, see Smith (1999)

Figure 3: Average Sick Days and Health over Income Quintiles



Notes: Dashed lines (right axis) show average self-reported health and 95% bootstrap confidence interval. Health is reported on ordinal five point scale where 0 denotes "bad" health and 4 denotes "very good" health. Solid lines (left axis) present average sick days of workers(40 – 50) and 95% bootstrap confidence interval separated for income quintiles.

Table 1: OLS and Logit Regressions of Sick Days on log Income

Sick Days	(1)	(2)	(3)
Log income	17.9951***	0.0828***	-2.124***
Log income squared	-1.1451***	—	—
Health	-1.7628***	-0.0867***	-1.5439***
Wave	-0.0867***	0.0015**	-0.1780***
Unemployment rate	-0.1583***	-0.0073***	-0.1627***
Observations	76,281	76,281	42,019

Notes: *** Significant at $\alpha = 0.01$, ** Significant at $\alpha = 0.05$, * Significant at $\alpha = 0.1$.

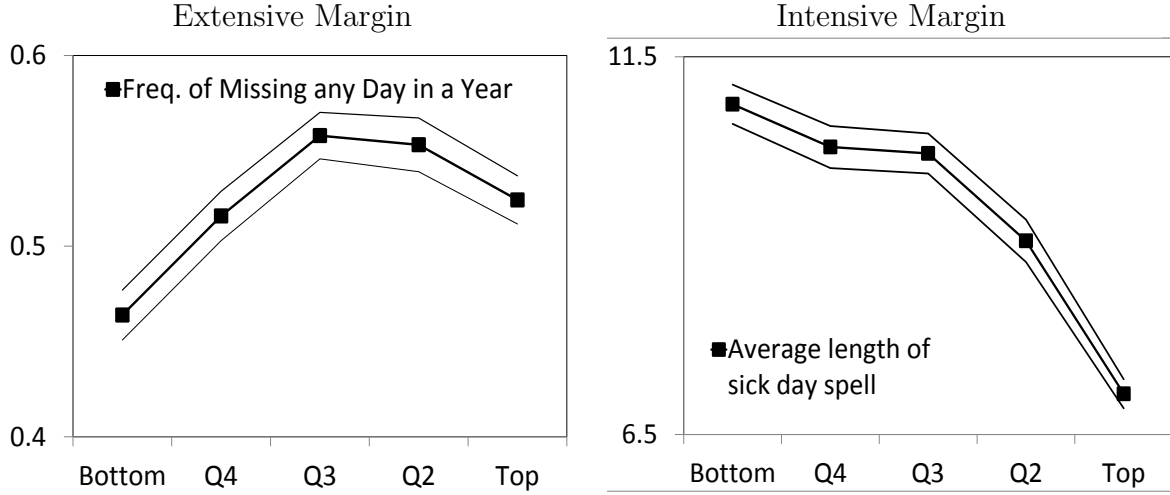
All regressions used the benchmark sample. Additionally to shown controls the regressions include sex, age, age² years of education and health². Robust standard errors are clustered on the personal level. Column (2) shows marginal effects at means.

Further insights to the characteristics of sick day use are provided by distinguishing between extensive margin, i.e., missing any day a year or not, and intensive margin, i.e., conditional on missing at least one day a year how many days the respondent is not at work.

The left panel of Figure 4 shows the extensive margin of sick days for workers (40 – 50). Workers in the bottom income quintile exhibit the lowest probability to miss any day in a year. The higher the income group the higher the probability to miss at least one day. Only at the very top the extensive margin seems to decrease. This pattern is also confirmed by estimating Equation (2) using the extensive margin as dependent variable. The results are presented as odd ratios in the second column of Table 1. The estimate for log income is highly significant and confirms the positive relationship between income

and a high probability to miss any day. Other results show an unsurprising protective effect of self-reported income against missing any day.

Figure 4: Composition of Sick Days over Income Quintiles



Notes: Left panel: Frequency of absence at least one day a year (extensive margin) at work of workers(40 – 50) and 95% bootstrap confidence interval separated for income quintiles. Right panel: Average number of sick days conditional on being at least one day sick (intensive margin) and 95% bootstrap confidence interval of workers(40 – 50).

The right panel of Figure 4 shows the intensive margin of sick days for workers (40 – 50). In contrast to the extensive margin the intensive margin is monotonically decreasing over income quintiles. The decline in conditional averages originates from different distribution length of sick day spells. The top income quintile have a higher probability of experience short sick spells (up to 14 days). On the other hand workers in the bottom income quintile have a high probability to have longer sick spells (more than 14 days).²⁷ The third column of Table 1 displays results of estimating Equation (1) using the intensive margin as dependent variable. The results confirm that a high income has a protective effect against sick days. The higher the income the fewer are the number of sick days a year conditional on being sick. Coefficients on health and recession have again the expected sign.

Summarizing, there is a remarkable difference in utilization of sick leave between income groups. The top income group has a low average but the highest probability to miss a day and the highest probability of a short spell when sick. The medium income quintile has the highest average of sick days. The bottom income quintile has the worst health but the lowest number of sick days on average. It also shows a huge discrepancy in terms of lengths of sick spells. It has the highest probability of missing not one days at work but the lowest probability of missing only few days and then again the highest probability of missing more than 14 days.

²⁷The density function of sick days for bottom and top income quintile is contained in Appendix 7.3.

2.3 Micro Evidence Using Panel Data

The panel structure of GSOEP allows to carve out further facts about sick days. First of all sick days are persistence. People that report sick days report to have also higher sick days in the next year. Including lagged sick days in estimating Equation (1) yields a positive and significant estimate. All other results remain qualitatively unchanged.

More important is the relation of sick days and unemployment. Table 2 shows descriptive statistics of workers that report employment in the previous year. This group is separated in two sub-samples. One in which respondents also work in the current period and one in which worker report to be unemployment in the current period. Workers who are laid off miss in the previous year on average 1.7 days more at work. I additionally compute the average of sick days of a worker over the last three years. The average sick days are almost one day higher for those that lost their job. Laid off workers are also poorer, less healthy and more likely to be male.

Table 2: Summary statistics of employed workers

	employed in prev. period	stayed employed	become unemployed
Sick days in prev. year	5.55	5.48	7.16
Av. sick days in prev. 3 years	6.78	5.29	6.28
Age	40.68	40.74	39.30
Income in prev. year	2,836€	2,881€	1,782€
Health in prev. year	2.62	2.64	2.57
Male	63%	63%	59%
Observations	109,559	104,897	4,662

Notes: Descriptive statistics for sample used in panel Logit model. Only workers that were employed in the previous year.

Table 3 shows the results of estimating Equation (3). The first column present the regression coefficients that later will be used in the structural model. The second column presents the results for the same regression but also including additional controls variables. e.g. sex and education. In both columns lagged income quintile has a negative sign, i.e., the highest probability of becoming unemployed is among workers in the bottom income quintile. Good health on the other side has a protective effect against the risk of unemployment. Healthier worker are less likely to become laid off. The key results are the coefficients of number of sick days on the risk to become unemployed. In both columns sick days show positive and highly significant results. The higher the number of sick days the more likely it is to become unemployed in the next year. ²⁸

²⁸In the relationship between sick days and layoffs might also exist the problem of reverse causality. Workers that know that they will lose their job could take sick days without fear of retaliation. This effect should be attenuated in the average sick days regression. I also run a regression for lagged sick days, i.e. number of sick days in the second to last year. They also show positive significant results. The

In a next step I show that these results also hold estimating Equation (4) including the fixed effect component. For a better understanding I report in the last two columns the results as odd ratios. The effect of sick days on unemployment is qualitatively not effected and still highly significant. The results for the effect of income and the recession dummy on risk of unemployment remain unchanged. Health has still the same qualitative sign but becomes insignificant in the model without fixed effects. In the last column I replace sick in the previous year for accumulated sick days over the last two years. The effect seems robust.²⁹

Table 3: Panel Results for Unemployment I

Unemployment	Coefficients		Odd Ratios	
	(1)	(2)	(3)	(4)
Sick Day prev. Year	0.0295***	0.0289***	1.0180***	—
Sick Day prev. 3 Years	—	—	—	1.0416***
Lagged Income Quintile	-0.5710***	-0.5890***	0.9276	1.0383
Lagged Health	-0.0862**	-0.0920**	0.9175	0.9423
Age	0.0136***	0.0153***	0.7468	0.8602
Recession	0.3792***	0.2849***	1.9604***	1.8195***
Controls	No	Yes	Yes	Yes
Fixed Effects	No	No	Yes	Yes

Notes: *** Significant at $\alpha = 0.01$, ** Significant at $\alpha = 0.05$, * Significant at $\alpha = 0.1$.

Controls include sex, years of education, and year dummies. Column (3) and (4) report odd ratios at population average. Robust standard error clustered.

3 Full Model

In this section I describe a dynamic stochastic model of work absence decisions. It captures both standard consumption-saving decision and sequential decision-making behavior of employed individuals regarding their health. I will later use this model to evaluate the consequences of economic inequalities for utilization of sick leave and health as well as the consequences emanating from different paid sick leave systems.

problem with this specification is that the sample size is heavily reduced. I exclude in this specification all respondents that were unemployed in the previous year. Higher number of sick days the second to last year however increased the likelihood of unemployment in the last year and therefore the exclusion.

²⁹In later sections it will be necessary to see whether the effect of sick days on unemployment is cyclical or not. To investigate this I estimate Equation (3) with fixed effects using an interaction term of *Recession* and the two measures of sick days. Both regressions show that the interaction term is positive but not significant. It seems that the influence on sick days on the probability of being laid off is the same in a recession as in a boom. Results are provided in the Appendix

3.1 Household's Problem

Agents start their economic life at age 20, retire at age 65. While retired the problem of the agent reduces to a consumption-saving decision. The agents live definitely to their terminal life of 80.³⁰

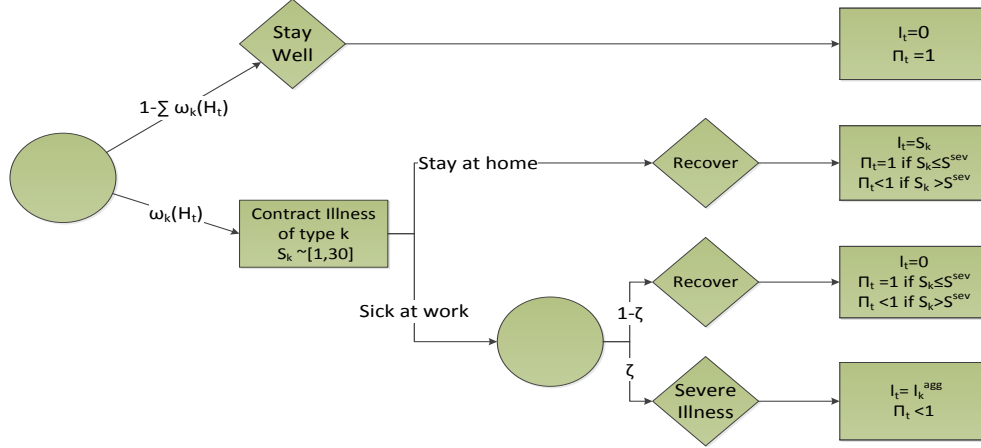
Health and Acute Illness The model features two distinct types of health: general health and acute sickness. General health, H_t , reflects the general constitution of an agent. It is a persistent state that only adjusts gradually. Agents start their economic life in a certain health state, H_0 . At the end of each period the agent faces a probability $1 - \Pi$ to drop into the lower next health state.

In contrast, acute sickness S_t has a transitory notion and should mimic the contraction of an illness, e.g., the flu. At the beginning of each period an individual faces the risk to either stay well, $S = S_0$, or contracting one of nine types of acute illnesses, $S = S_k$. These acute illnesses vary in their severity, i.e., they differ in the time required to recover as well as the dis-utility they incur. The probability of contracting an illness of type k , $\omega_k(H_t)$, depends only on the overall health status H_t of an individual. The lower the general health state the higher is the probability to contract a severe illness. The sickness type affects the number of sick days and the probability of dropping in health $1 - \Pi_k$.

When staying well, $S = 0$, there is no need to take sick days and also no chance of a lower health state in the next period. Upon becoming ill, individuals can decide whether or not to be absent from work. By staying home from work, the individual is required to take $l_t = S_k$ sick days according to the severity of the illness. The probability to drop in health then depends on the severity of sickness. For all sickness states that require more than S^{sev} sick days there is a positive probability to drop in health state. By going sick to work, the worker faces two possible outcomes. With probability $1 - \zeta$ the worker recovers without taking recuperation time, i.e., $l_t = 0$. With probability ζ the sickness aggravates. In this case the worker is forced to stay at home for $l_t = l_k^{agg}$ sick days. Additionally the individual faces the probability to drop in her health state, independent of the initial sickness state. Note that the decision -about going to work sick or not- is made simultaneously to the saving consumption decision at the beginning of the period. Workers cannot change their consumption in the case their sickness aggravates. The model also ignores preventive treatment, so individuals cannot improve their overall health by taking sick days without being sick.

³⁰Unlike most of the health economics literature I abstract from survival rates.

Figure 5: Events and Decisions - Acute Illness



Notes: Circles denote events, rectangles denote decisions, diamonds denote outcome.

Preferences Individuals value consumption and dislike being sick over the life cycle c_t, S_t according to a standard time-separable utility function

$$\mathbb{E} \left\{ \sum_{j=0}^J \beta^j u(c_t, S_t) \right\}.$$

where β is the raw time discount factor and expectations are taken over stochastic employment and health history.

Following Finkelstein et al. (2013) I assume that the period utility function is given by

$$u(c_t, S_t) = (1 + \psi S_t) \frac{c_t^{1-\sigma}}{1-\sigma}.$$

It is a standard CRRA utility function for non-sick individuals. Sickness multiplies the marginal utility of period consumption by a factor of $(1 + \phi)$.³¹ Noting that there are situations where health and consumption are substitutes (e.g., marginal utility of hiring a maid is higher for a sick person) I choose to model consumption and sickness complements (e.g., marginal utility of a ski trip is lower for a sick person). The parameter σ determines the inter-temporal elasticity of substitution.

Employment Status and Income Agents differ with respect to their age and their predetermined type. These sources of heterogeneity affect an agent's labor productivity which is given by $\Gamma_{k,j}$. First, the labor productivity differs according to the age of an agent j . Second, each household belongs to a particular group k that shares the same average productivity. Differences in groups stand in for differences in education or ability,

³¹Sickness and health may also have a direct effect on the level of utility. I abstract from this in the benchmark calibration.

characteristics that are fixed at entry into the labor market and affect a groups relative income. I introduce these differences in order to generate part of the cross-sectional income, cf. Krueger and Ludwig (2007). Notice that health does not directly affect labor productivity. The wage rate is w_t .

$$y_{t,k,j} = \Gamma_{k,j} w_t.$$

Besides labor productivity, income of an agent depends crucially on the employment status I . At the end of each period the agent may be dismissed. Central to this paper is that workers can reduce the layoff probability in the next period by reducing the number of sick days. Theoretically, there exist different explanations for this linkage. First, health is important for the productivity of a worker. Employers cannot directly observe health and use sick days as a signal for health. Table 1 shows that this is true and sick days are driven by health. Second, related to the shirking literature, employer can only imperfectly monitor workers in terms of sick days. The higher the number of sick days the higher is the probability to be discovered. Third, due to the structure of the German system, sick days are costly for the employer and persistent over time.³² Therefore employer's incentive is high to get rid of worker with high sick days to save these costs. In all these cases higher sick days lead to a higher layoff probability. For the following structural model it is secondary why sick days are a good predictor of future unemployment. It is important that worker know about this fact and take it into account when optimizing their number of sick days.³³ The probability to keep the job, ϕ , depends additionally on age, skill type, health of the worker as well as the current unemployment rate.

$$\phi = \phi(k, j, H, \bar{u}, l)$$

While unemployed the agent gets unemployment benefits, $b_{j,k}^U$, which depend on age and type of the agent. The probability to find a new job when unemployed, $\bar{\phi}$, again is determined by age, skill health, and general economic conditions. They ,however, do not depend on the number of sick days.

$$\bar{\phi} = \bar{\phi}(k, j, H, \bar{u})$$

The agent gets only paid for the time she works $1-l_t$. For the sick time l_t the household could get reimbursed by the government with payments depending on the regular labor income $b_{j,k}^S$.³⁴

³²High number of sick days are followed by high number of sick days. A simple test for auto-correlation yields a coefficient of 0.34.

³³It would be nice to do that for the accumulated sick days over some periods but this would require an additional endogenous state and is computational burdensome.

³⁴This reimbursement is actually not done by the government but by the employer. As I do not model

The evolution of the unemployment rate in my model is exogenous (i.e., I do not model general equilibrium effects or the firm side) and is the main driving force behind the model. For simplicity I model only two states of the general economic conditions, i.e., "boom" and "recession".

Budget constraint and borrowing limit Individuals can accumulate assets, a , at a constant interest rate R . They are not allowed to borrow. They allocate their total resources between consumption c , and asset holdings for next period:

$$\begin{aligned} (1 - \tau_t)(1 - l_t)y_{k,j,t} + l_t b_{j,k}^S I_t + b_{j,k}^U (1 - I_t) + Ra_t &= a_{t+1} + c_t & \text{for } j \leq 65 \\ b_{j,k}^R + Ra_t &= a_{t+1} + c_t & \text{for } j > 65 \end{aligned}$$

3.2 Government Policies

The government imposes a flat income tax, τ . The collected revenues are used for three main purposes: (i) to finance the unemployment insurance b^U , (ii) to finance the paid sick leave b^S and (iii) finally, to finance the pension system b^R . I assume that the budget of the government is balanced at all times.

$$\begin{aligned} \sum_i \sum_{j=0}^{45} \sum_k \tau_t y_{t,i,j} (1 - l_{t,i,j}) I_{t,i,j} = \\ \sum_i \sum_{j=0}^{45} \sum_k [b_{j,k}^U (1 - I_{t,i,j}) + b_{j,k}^S I_{t,i,j} l_{t,i,j}] + \sum_i \sum_{j=46}^{60} \sum_k [b_{j,k}^R] \quad (5) \end{aligned}$$

3.3 Individual's Dynamic Program

I model the decisions to miss work during an episode of acute illness as the sequential choices of workers solving a discrete choice stochastic dynamic programming problem. At each discrete period of an illness the forward-looking individual chooses whether or not to miss work based on expected utility maximization.

Individuals, at the beginning of period t are indexed by their age j , their group k , their asset holdings a , their health state H , their job employment status I and their realization of acute sickness, S . To simplify the analysis, I assume that the factor prices are exogenous. Each individual starts their life in a specific health state H_0 and is endowed

the firm site I make this shortcut. However, it would be interesting to see how this different setting would alter the result in a general equilibrium.

with initial assets a_0 . Thus their maximization problem reads as

$$W(j, k, a_t, H_t, I_t, S_t) = \max_{c_t, l_t, a_{t+1}} u(c_t, S_t) + \beta \sum_{H_{t+1}} \sum_{I_{t+1}} \Pi(H_{t+1} | H_t) \Phi(I_{t+1} | I_t) \sum_{S_{t+1}} \omega(H_{t+1}) W(j+1, k, a_{t+1}, H_{t+1}, I_{t+1}, S_{t+1}) \quad (6)$$

subject to the constraints

$$\begin{aligned} a_{t+1} + c_t &= (1 - \tau_t)(1 - l_t) y_{k,j,t} + l_t b_{j,k}^S I_t + b_{j,k}^U (1 - I_t) + R a_t \\ a_{t+1} &\geq 0 \\ I_{t+1} &= \begin{cases} 1 & \text{with probability } \Phi_t \\ 0 & \text{with probability } 1 - \Phi_t \end{cases} \\ \Phi_t &= \begin{cases} \phi(k, j, h_t, \bar{u}_t, l_t), & \text{if } I_t = 1 \\ \bar{\phi}(k, j, h_t, \bar{u}_t), & \text{if } I_t = 0 \end{cases} \\ \Pi_t &= \Pi(H_t, S_t, l_t) \end{aligned}$$

Definition 1 *A stationary competitive equilibrium of this economy for given sick pay schemes b^S , unemployment benefits b^U , retirement benefits b^R , tax rate τ , wage w , and risk-free interest rate r is a set of decision rules, $\{c_t(z), l_t(z), a_{t+1}(z)$ and value functions W where $z = (t, j, k, a_t, H_t, I_t, S_t)$ such that:*

1. *Given initial conditions $W(t, \cdot)$ solves Eq. (6) and decision $c(t, \cdot)$, $l(t, \cdot)$, $a_{t+1}(t, \cdot)$ are the associated policy functions.*
2. *Government policies satisfy Eq.(5) in every period.*

4 Quantitative Analyses

In this section, I begin by discussing the parameter choices for the model. Then in Section 4.2 I present simulation results and their counterparts in the data to evaluate the model's performance such as the lifetime profile of sick days by income, health differences over the life cycle, etc.

4.1 Parameter Estimation and Calibration

In this section I discuss the specification of the model parameters. I need to choose parameters governing the employment status, health transitions, preferences, and policy settings. The determination of the model parameters proceeds in three steps. First, I fix a subset of parameters exogenously. Second, parts of the model parameters can be estimated from the GSOEP data directly. These include the parameters governing the

probability to get laid off ϕ , find a new job, $\bar{\phi}$, the probability to contract an illness conditional on the health state ω ; as well as the productivity difference $\Gamma_{k,j}$. Third, (and given the parameters obtained in step 1 and step 2) the remaining parameters (mainly those governing the health transition Π , ζ , and health preferences ψ) are then determined through a method of moments estimation of the model with GSOEP data. I now describe these three steps in greater detail.

4.1.1 A Priori Chosen Parameters

Table 7.4 show the parameter that are fixed exogenously with their values. The model period is one year. The life span of an individual is $J = 45$ periods. I assume that the interest rate, R is determined exogenously by world factors in an open-economy equilibrium and following Fernandez-Villaverde and Krueger (2011) I set $R = 4\%$. Then I select two preference parameters. Consistent with values commonly used in the quantitative macroeconomics literature I choose a risk aversion parameter of $\sigma = 2$ and a time discount factor of $\beta = 0.96$ per annum. I choose $\sigma = 2$ to obtain an inter-temporal elasticity of substitution of 0.5, which is a value widely used in the literature (e.g. Fernandez-Villaverde and Krueger (2007)).

Policy Parameters For the benchmark calibration I choose the current institutional setting for Germany. It shuts down two direct effects of income on health. First Germany has universal health care coverage. So individuals do not have to pay for standard medical expenditures, e.g. doctor visits. The second I set the benchmark paid sick leave coverage to 100% of the current wage. So there is no direct reduction in income of sick leave for individuals. This is important to isolate the indirect effect of income on health via risk of unemployment. I set the unemployment benefits to 60% of the former wage of a worker. This is the current German setting for the first 12-24 month in Germany.³⁵ I set the retirement benefits to 50% of the former wage of a worker.

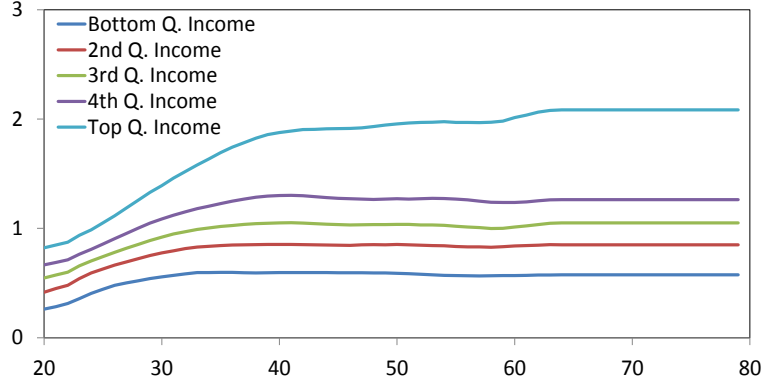
4.1.2 Parameters Estimated Directly from the Data

In a second step I estimate part of the model parameters directly from the data, without having to rely on the equilibrium of the model. Results are shown in Table 9 - 11 in Appendix 7.4.

Labor Productivity Using the GSOEP data on income I compute the age-dependent productivity for five different income types. Each individual therefore has a certain productivity of $\Gamma_{k,j}$. Figure 6 show the age dependent income profiles.

³⁵The unemployment setting has undergone a major reform in 2005. This might limit the historical comparison between model output and GSOEP data.

Figure 6: Income Profile for Income Quintiles over Life Cycle



Notes: Income age profiles for different skill groups.

Job Keeping Probability The probability to lose the job $1 - \phi$ is directly taken from the estimation in section 2.3. The model computes for each combination of skill k , age j , health status h , economic state $\bar{u}$ and number of sick days l the predicted probability to retain the job. The used coefficients are taken from the first column of Table 3.

$$\hat{y} = \alpha_{0,1} + \alpha_{1,1} * k + \alpha_{2,1} * j + \alpha_{3,1} * H + \alpha_{4,1} * \bar{u} + \alpha_{5,1} * l$$

$$1 - \phi(k, j, H, \bar{u}, l) = 1 - \frac{e^{\hat{y}}}{1 + e^{\hat{y}}}$$

The probability to stay unemployed $\bar{\phi}$ is computed the same way except using sick days.³⁶

$$\hat{y} = \alpha_{0,2} + \alpha_{1,2} * k + \alpha_{2,2} * j + \alpha_{3,2} * H + \alpha_{4,2} * \bar{u}$$

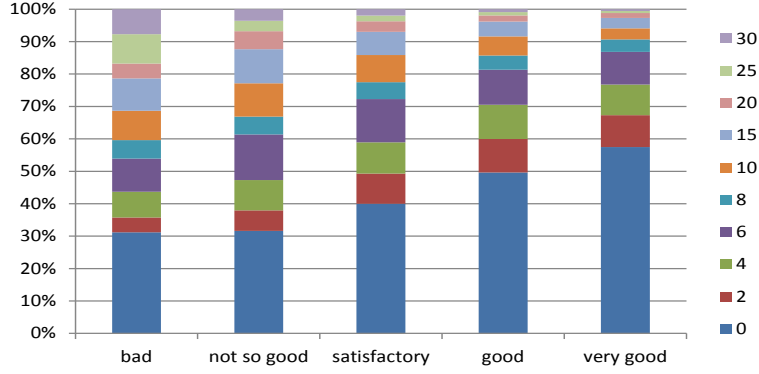
$$1 - \bar{\phi}(k, j, H, \bar{u}) = 1 - \frac{e^{\hat{y}}}{1 + e^{\hat{y}}}$$

Incidence of Acute Sickness The incident of an acute illness depends on the current health state of the agent. Individual with low health have both a higher probability to contract an illness and a higher probability that illness is more severe. The according probabilities ω I take from the GSOEP. For computational reason I restrict the number of sick days to nine sick day bins, [1-2, 3-4, 5-6, 7-8, 9-10, 11-15, 16-20, 21-25, 26-30]. For each of these bins I compute the frequency conditional on the health status and being in the top income quintile.³⁷ Note that I make the implicit assumption that the top income quintile always take the sick days that are required to recover. Figure 7 show the frequency conditional on health.

³⁶The corresponding regression results are contained in Appendix Table 3.

³⁷Note that the probability of contracting an illness as well as the health transition do not depend on age. In this I am in line with Dalgaard and Strulik (2014) that show that health do not depend on the calendrical age but is an accumulation of deficits.

Figure 7: Frequency of Sick Days for Health States (Top Income Quintile)



Notes: Frequency of sick days distributed over sick day bins. Conditional on top income quintile but not age.

4.1.3 Parameters Calibrated Within the Model

In a final step I now use my model to find parameters governing the transition probabilities for health, the probability that sickness aggravates, and preference parameters governing utility from health. All estimated parameters can be found in Table 13 in Appendix 7.4.

Health Transition Empirically the health stock, H_t , is defined using self-reported health status. Thus H_t takes one of five values: 1-”poor”, 2-”fair”, 3-”good”, 4-”very good”, 5-”excellent” that are contained in the GSOEP.

In the benchmark case I calibrate the initial health state distribution, H_0 , and the health transition probabilities, Π such that the distribution of health states of the top income quintile match the data at the beginning and the end of working life.

Acute Sickness Important for the decision to stay at home to recover or go sick to work is the aggravation probability, $0 < \zeta < 1$, and the number of sick days l^{sev} required when the sickness aggravated. I assume the following relationship holds

$$\mathbb{E}[l_t | S_k, \text{Sick to work}] = \kappa S_k$$

where $0 < \kappa < 1$. This leads to the relationship

$$l^{sev} = \frac{S_k}{\kappa \zeta}$$

I calibrate this to match the average number of sick days.

Health Preference The parameter ψ that determines the relative importance of sickness in the utility function. It is chosen such that top income quintile never go sick to

work. Borrowing constraint households however cannot follow this optimal strategy and will go sick to work.

Tax I adjust the tax rate τ to have a balanced per period budget, e.g. to finance paid sick leave, unemployment and retirement benefits.

4.2 Model Fit and Benchmark Results

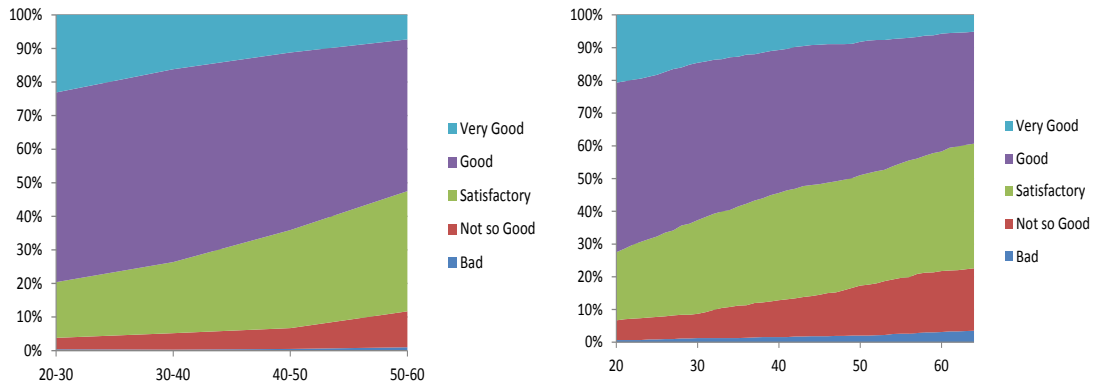
In this section, I examine the fit of the model to the data. First I discuss the performance of the model in fitting the targeted moments in the estimation. Then I present an informal over-identification test of the model by showing the model's performance in fitting the moments that are not targeted in the estimation.

In order to do so I simulate 100,000 life-cycle paths for individuals. For each life cycle path the initial states for health, H_0 , and the employment status, I_0 , are drawn from distribution that match the data. All individuals start with no initial assets, $a_0 = 0$.

4.2.1 Fit of the Model to the Targeted Moments

The upper and the lower left panel of Figure 8 exhibits the distribution of health states of individuals for the top and bottom income quintile, respectively. The right panels shows the data counterpart estimated from the GSOEP. The model is able to account for the key health profiles over the life cycle. They also exhibit the differences between the poor and the rich. Note that both income groups start with the same initial health distribution at age 20.

Figure 8: Health States over the Life Cycle for Top Income Quintile

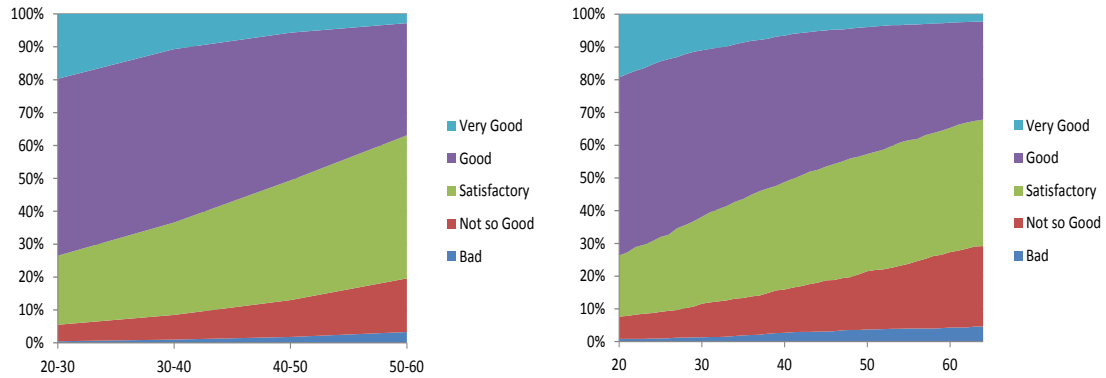


Notes: Upper Left: Data bottom income quintile. Upper Right: Model outcome bottom income quintile. Lower Left: Data top income quintile. Lower Right: Model outcome top income quintile.

4.2.2 Fit of the Model to Non-Targeted Moments

One moment that is not targeted in the calibration is the health profile over the life cycle of the poor income quintile. Figure 9 shows the distribution of sick state for the data and the model.

Figure 9: Health States over the Life Cycle for Top Income Quintile



Notes: Upper Left: Data bottom income quintile. Upper Right: Model outcome bottom income quintile. Lower Left: Data top income quintile. Lower Right: Model outcome top income quintile.

[To be added]

- average cross section sick days
- intensive and extensive margin

5 Policy Evaluation

The determination of policy-invariant structural parameters allows for the introduction and evaluation of different policies that affect the financial constraints of a consumer's decision-making problem. The policy instrument in the paper involves sick leave coverage.

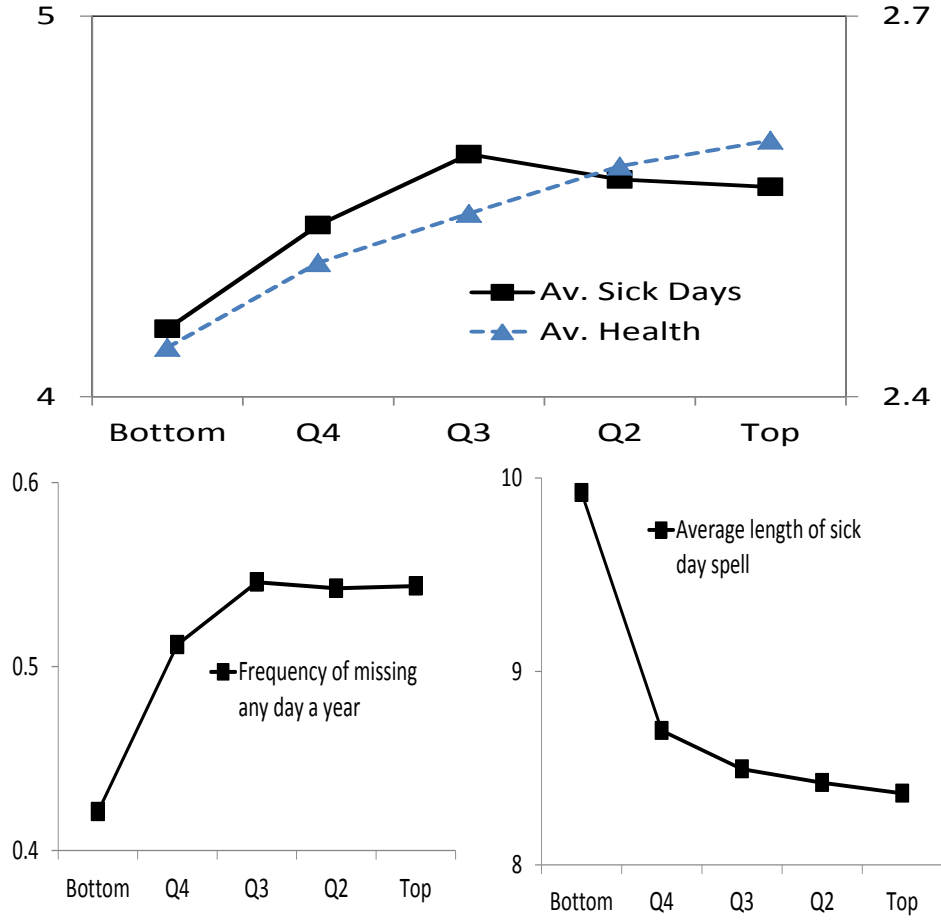
[To be added]

US - No mandatory SP

UK - Wage independent SP

France - Grace period for SP

Figure 10: Average Sick Days and Health over Income Quintiles



Notes: Dashed lines (right axis) show average self-reported health and 95% bootstrap confidence interval. Health is reported on ordinal five point scale where 0 denotes "bad" health and 4 denotes "very good" health. Solid lines (left axis) present average sick days of workers(40 – 50) and 95% bootstrap confidence interval separated for income quintiles.

6 Conclusion

In this paper, I studied the relationship between sick leave and unemployment and its effect on the evolution of health. Using data from the GSOEP I document new empirical facts on sick leave by income quintiles. First, the average number of sick leave exhibits a hump shaped pattern over income quintiles. Second, bottom and top income worker differ significantly in the pattern of sick leave. Poor individuals try to not miss any day or take a high number of sick days. Rich individuals take constantly a moderate number of sick days.

I develop and estimate a life cycle model of health that can account for these facts. The main feature of my model is entangle the employment risk with health-related decisions. Moreover, I incorporate important features of the German labor market system into my model, such as universal health care and sick leave coverage.

I estimate my model using both micro and macro data. Then I use my model to

analyze the macroeconomic effects of a counterfactual policy analysis.

References

- Ales, L., R. Hosseini, and L. E. Jones (2012, March). Is There “Too Much” Inequality in Health Spending Across Income Groups? NBER Working Papers 17937, National Bureau of Economic Research, Inc.
- Andersen, S. H. (2010). The cost of sickness: On the effect of the duration of sick leave on post-sick leave earnings. *Social Science & Medicine* 70(10), 1581–1589.
- Arai, M. and P. S. Thoursie (2005). Incentives and selection in cyclical absenteeism. *Labour Economics* 12(2), 269–280.
- Askildsen, J. E., E. Bratberg, and ivind Anti Nilsen (2005). Unemployment, labor force composition and sickness absence: a panel data study. *Health Economics* 14(11), 1087–1101.
- Bergström, G., L. Bodin, J. Hagberg, T. Lindh, G. Aronsson, and M. Josephson (2009). Does sickness presenteeism have an impact on future general health? *International Archives of Occupational and Environmental Health* 82(10), 1179–1190.
- Bundeministerium für Arbeit und Soziales (2014). Sozialbudget 2013. Technical report, Bundesministerium für Arbeit und Soziales.
- Bundesagentur für Arbeit (2014). Arbeitslosigkeit im zeitverlauf. Technical report, Bundesagentur für Arbeit.
- Cole, H. L., S. Kim, and D. Krueger (2014, April). Analyzing the Effects of Insuring Health Risks: On the Trade-off between Short Run Insurance Benefits vs. Long Run Incentive Costs. PIER Working Paper Archive 14-023, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.
- Dalgaard, C.-J. and H. Strulik (2014). Optimal aging and death: Understanding the preston curve. *Journal of the European Economic Association*, n/a–n/a.
- Ehrlich, I. and H. Chuma (1990, August). A Model of the Demand for Longevity and the Value of Life Extension. *Journal of Political Economy* 98(4), 761–82.
- Fernndez-Villaverde, J. and D. Krueger (2007, August). Consumption over the Life Cycle: Facts from Consumer Expenditure Survey Data. *The Review of Economics and Statistics* 89(3), 552–565.
- Fernndez-Villaverde, J. and D. Krueger (2011, November). Consumption And Saving Over The Life Cycle: How Important Are Consumer Durables? *Macroeconomic Dynamics* 15(05), 725–770.
- Finkelstein, A., E. F. P. Luttmer, and M. J. Notowidigdo (2013). What good is wealth without health? the effect of health on the marginal utility of consumption. *Journal of the European Economic Association* 11, 221–258.

- Gilleskie, D. B. (1998, January). A dynamic stochastic model of medical care use and work absence. *Econometrica* 66(1), 1–46.
- Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy* 80(2), 223–55.
- Hall, R. E. and C. I. Jones (2007, 02). The Value of Life and the Rise in Health Spending. *The Quarterly Journal of Economics* 122(1), 39–72.
- Halliday, T., H. He, and H. Zhang (2012, June). Health investment over the life-cycle. Working Papers 201210, University of Hawaii at Manoa, Department of Economics.
- Hesselius, P. (2007). Does sickness absence increase the risk of unemployment? *The Journal of Socio-Economics* 36(2), 288–310.
- Kivimäki, M., J. Head, J. E. Ferrie, H. Hemingway, M. Shipley, J. Vahtera, and M. G. Marmot (2005). Working while ill as a risk factor for serious coronary events: The whitehall ii study. *American Journal of Public Health* 95(1), 98102.
- Krueger, D. and A. Ludwig (2007, January). On the consequences of demographic change for rates of returns to capital, and the distribution of wealth and welfare. *Journal of Monetary Economics* 54(1), 49–87.
- Leigh, J. P. (1985). The effects of unemployment and the business cycle on absenteeism. *Journal of Economics and Business* 37(2), 159–170.
- Markussen, S. (2012, October). The individual cost of sick leave. *Journal of Population Economics* 25(4), 1287–1306.
- Nuebling, M., H. H. Andersen, A. Muehlbacher, J. Schupp, and G. G. Wagner (2007). Computation of Standard Values for Physical and Mental Health Scale Scores Using the SOEP Version of SF12v2. *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift fr Wirtschafts- und Sozialwissenschaften* 127(1), 171–182.
- Ozkan, S. (2011). Income Differences and Health Care Expenditures over the Life Cycle. 2011 Meeting Papers 478, Society for Economic Dynamics.
- Pfeifer, C. (2013). Cyclical absenteeism among private sector, public sector and self-employed workers. *Health Economics* 22(3), 366–370.
- Puhani, P. A. and K. Sonderhof (2010). The effects of a sick pay reform on absence and on health-related outcomes. *Journal of Health Economics* 29(2), 285–302.
- Ruhm, C. J. (2000). Are recessions good for your health? *The Quarterly Journal of Economics* 115(2), 617–650.
- Scheil-Adlung, X. and L. Sandner (2010). The case for paid sick leave. World Health Report Background Paper 9, World Health Organization.

- Smith, J. P. (1999). Healthy bodies and thick wallets: The dual relation between health and economic status. *The Journal of Economic Perspectives* 13(2), pp. 145–166.
- Wagner, G. G., J. R. Frick, and J. Schupp (2007). The german socio-economic panel study (soep) scope, evolution and enhancements. *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift fr Wirtschafts- und Sozialwissenschaften* 127(1), 139–169.
- Ziebarth, N. R. (2013). Long-term absenteeism and moral hazardEvidence from a natural experiment. *Labour Economics* 24(C), 277–292.
- Ziebarth, N. R. and M. Karlsson (2010). A natural experiment on sick pay cuts, sickness absence, and labor costs. *Journal of Public Economics* 94(11-12), 1108–1122.

7 Appendix

7.1 German sick leave policy

Compulsory sick pay with 100% wage replacement was established 1930 for white collar employees and 1969 for blue collar workers. The current regulation of sick leave (*Entgeltfortzahlung im Krankheitsfall*) in Germany is determined in the *Entgeltfortzahlungsgesetz*. According to the law eligible for paid sick leave are those employees (also including part time and temporary workers) that fulfill following conditions:

- The employment has to be in place for four weeks.
- The worker has to be incapable of working.
- The incapability has to be a consequence of an illness.
- The illness is not a result of a gross negligence.

The sick pay has to be provided by the employer for the length of six weeks. If a worker becomes sick again with the same sickness then the sick days are summed up until the six weeks are reached. The claim of sick pay renews if the worker suffers from a different illness or it has been more than 6 months that the worker was sick with the same illness. The worker receives the wage that she would have earned if she hasn't got sick. There is no grace period. If a worker becomes sick while she is on vacation her holiday entitlement is not reduced. The worker has to tell her employer immediately about the incapability of work. On the fourth day of the sick spell the worker has to send a sick certificate issued by practitioner.

Monitoring the worker is highly restricted. The German Federal Labor Court decided that the observation of an employee by her employer is illegal without concrete evidence supporting the suspicion of fraud. (Court decision 19. February 2015 - 8 AZR 1007/13). The employer can only request that the employee is reexamined by the practitioner of the Medical Service of the Health Funds (Medizinischer Dienst der Krankenversicherung).

Between October 1996 and December 1998 there was a temporary change in the law. The main changes were a reduction of wage replacement from the 100% to 80%. However this reduction only applied to a fraction of the German work force as collective labor agreements between unions and firms mostly kept the 100% wage replacement. Empirical research on this law discontinuity is done by Ziebarth and Karlsson (2010) and Ziebarth (2013).

7.2 Sample selection

Table 4 shows the descriptive statistics before and after sample selection. I use in these statistics weighted samples. Weights are provided by the GSOEP to match the German

micro-census. The final sample is younger due to focusing on working age population. The higher percentage of men in the sample can be explained by their higher participation rate in the labor force. The average number of reported sick days is slightly increased after sample selection. The higher unemployment rate is due to the exclusion of the part time and temporary worker.

Table 4: Descriptive Statistics for Sample Selection

	Whole Sample	Benchmark Sample
Male	46%	62%
Age	49	41
Years of Education	11.7	11.9
Health	2.32	2.49
Income	2,355€	2,774€
Unemployed	6.79%	15.7
Sick	9.67	10.46
Observations	366,414	145,156

Notes: Descriptive statistics before and after sample selection. Benchmark sample used in the cross section and panel analysis.

7.3 Robustness Check in Empirical Part

Different measures of sick days

Sick days of worker have a skewed distribution. Table 5 provides results for the correlation of the unemployment rate and different measures of sick days. First, it shows the results for the median worker. The second column shows the correlation with the extensive margin, i.e., whether the respondent has missed a day or more or not. In the last three columns different cut of level for the maximum sick days are used. All of results are negative and in the same range as the benchmark result. The pro-cyclical pattern is extremely robust.

Table 5: Sick days and unemployment - Different sick day measures

Correlation	median	ext. marg.	max 120	max 60	max 30
Unemployment rate	-0.6405	-0.6329	-0.7466	-0.7601	-0.7282

Notes: Times series correlation of different measures of sick days and unemployment rate. First the sick days of the median respondent, second the cyclical behavior of the extensive margin. The last three columns represent the correlation of the mean with different cut off levels for maximum sick days.

Composition Effects

A potential different explanation of the cyclical behavior of sick days could arise if sectors (e.g., construction sector) with high usual high number of sick days are more prone to business cycles than the rest of the economy. To control that the general effect is not driven by this reason I check for different sector whether their exclusion alter the general finding. Table 6 shows the exclusion of the construction sector does not alter the benchmark result. The correlation coefficient is only slightly reduced to -0.71 .

I also check whether this cyclical behavior is different for different occupation type. GSOEP provides the ISCO88 classification and using the white/blue collar distinction as in the European working conditions surveys. Table 6 shows that for both subgroups the pro-cyclical behavior of sick days holds.

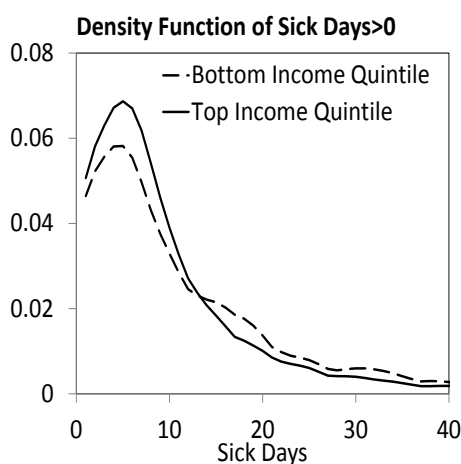
Table 6: Sick days and unemployment - Different sectors and occupations

Correlation	Without construction	Blue collar	White collar	Never unemployed
Unemployment rate	-0.7144	-0.6582	-0.6436	-0.6636

Notes: Times series correlation of average sick days and unemployment rate for different subgroups.

Density Function of Sick Days

Figure 11: Density function of Sick Days for bottom and top income quintiles



Notes: Density function of Sick days for bottom and top income quintile Age: 40 – 50

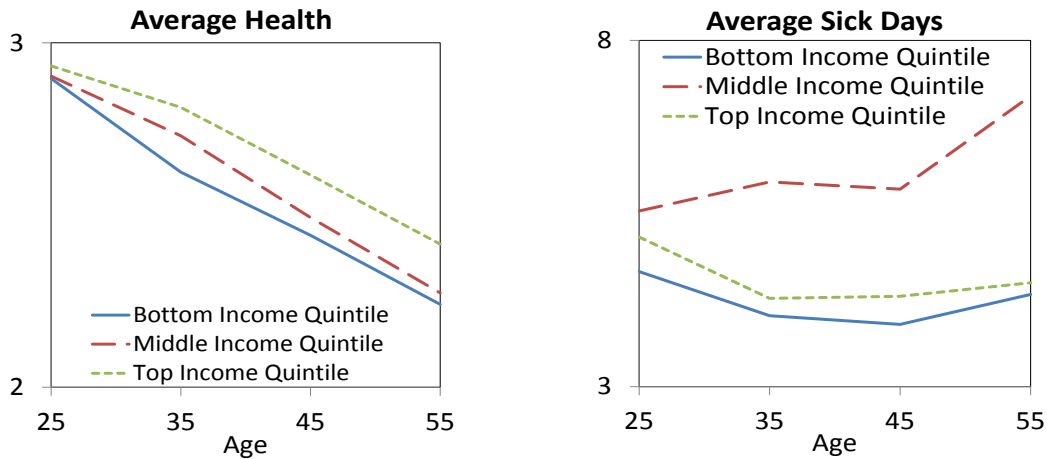
Age Profiles of Sick Days and Health

The right panel of Figure 12 confirms that the observed hump shaped income profile hold for all age-groups. The left panel also reveals that the income gradient in health

is increasing over age. The difference in health between bottom and top income quintile almost five times as high for 50 – 60 year old workers than for 20 – 30 year old.

Needs to be reviewed. Control for cohort effects.

Figure 12: Sick Days and Health over Life Cycle by Income Quintiles



Notes: Left Panel: Average self-reported health of bottom (solid), medium (dashed) and top income (dotted) quintile over the life cycle. Right Panel: Average sick days of bottom (solid), medium (dashed) and top income (dotted) quintile over the life cycle. Age bins: 20 – 30, 30 – 40, 40 – 50, 50 – 60

Controlling for gender in sick day profiles over income quintiles

still to be added

Interaction Recession Sick Days

Table 7 show the relation between unemployment, unemployment rate and sick days.

Table 7: Panel Results for Unemployment II

Unemployment	(1)	(2)
Sick Day prev. Year	1.0206***	—
Sick Day prev. 3 Years	—	1.0579***
Recession	1.6159***	1.5475***
Interaction	1.0052	1.0055
Fixed Effects	Yes	Yes

Notes: *** Significant at $\alpha = 0.01$, ** Significant at $\alpha = 0.05$, * Significant at $\alpha = 0.1$. All regressions include age, income, health and year dummies. Reported are odd ratios at population average. Robust standard error clustered.

7.4 Estimation Results

7.4.1 Fixed Parameters

Table 8: Fixed Parameters

Parameter	Description	Value
<i>Economy</i>		
T	Life span	45
R	Interest rate	0.04
w	Wage rate	1
<i>Preferences</i>		
β	Discount factor	0.9659
σ	Inter-temporal elasticity of substitution	2
<i>Policy</i>		
b^U	Unemployment benefit	60%
b^S	Sick leave replacement rate	100%
b^R	Retirement benefit	50%

Notes: Parameters taken from the literature.

7.4.2 Directly Estimated Parameters

Table 9: Coefficients for the layoff-probability

	Const.	K	J	H	$\bar{u}$	l
Employed	-4.2405	-0.5709	0.0136	-0.0862	0.0	0.0295
Unemployed	-0.0510	-0.0101	0.0134	-0.1213	0.0	-

Notes: Parameters estimated directly from GSOEP.

Table 10: Probability to become sick cond. health - ω

	S_0	S_2	S_4	S_6	S_8	S_{10}	S_{15}	S_{20}	S_{25}	S_{30}
H_1	0.3001	0.0277	0.0833	0.1250	0.0555	0.0833	0.1450	0.0555	0.0694	0.0555
H_2	0.3051	0.0641	0.0974	0.1435	0.0555	0.1025	0.1085	0.0538	0.0316	0.0376
H_3	0.3921	0.0961	0.0975	0.1350	0.0547	0.0840	0.0725	0.0312	0.0172	0.0193
H_4	0.4936	0.1047	0.1055	0.1086	0.0444	0.0590	0.0461	0.0185	0.0098	0.0093
H_5	0.5654	0.0967	0.0999	0.1030	0.0385	0.0340	0.0322	0.0179	0.0076	0.0044

Notes: Parameters directly estimated from GSOEP.

Table 11: Labor productivity over income groups and age - Γ

	K_1	K_2	K_3	K_4	K_5
J_{20}	0.224826037	0.328405423	0.456578007	0.587303446	0.732877732
J_{21}	0.250005566	0.409269009	0.560519145	0.685354277	0.850657669
J_{22}	0.311148441	0.509774966	0.622307323	0.724721362	0.884406497
J_{23}	0.350360594	0.552927485	0.656849717	0.754174695	0.924644522
J_{24}	0.425638928	0.598862981	0.700761760	0.809373763	0.979296359
J_{25}	0.453304887	0.633908213	0.749711757	0.848907970	1.043437128
J_{26}	0.486555339	0.666411652	0.784269193	0.909526063	1.105854970
J_{27}	0.505469995	0.689869703	0.819220306	0.957395861	1.196744849
J_{28}	0.526567728	0.729833055	0.848928302	1.009396469	1.256001953
J_{29}	0.534257324	0.749711757	0.892406923	1.043371438	1.333614431
J_{30}	0.552043113	0.775403025	0.923040554	1.094749559	1.398805958
J_{31}	0.588814584	0.811998956	0.964703602	1.141444947	1.461369255
J_{32}	0.581567748	0.813786468	0.967081217	1.143405056	1.507091996
J_{33}	0.590600670	0.827264526	0.997438473	1.182487624	1.598022908
J_{34}	0.610923651	0.844675221	1.0	1.195602273	1.638440613
J_{35}	0.605390496	0.850657669	1.019549932	1.236543287	1.696821211
J_{36}	0.587309058	0.841163678	1.035278366	1.258038975	1.730506925
J_{37}	0.595460368	0.850998814	1.034804554	1.262927700	1.803819125
J_{38}	0.588814584	0.851471889	1.046094062	1.291190800	1.845452180
J_{39}	0.598372424	0.854338956	1.050974415	1.293505670	1.844733642
J_{40}	0.596219571	0.856083687	1.053135456	1.324056682	1.900608466
J_{41}	0.596219571	0.850717011	1.053102151	1.303667321	1.891957676
J_{42}	0.598372424	0.854338956	1.051454666	1.293505670	1.899733341
J_{43}	0.587912961	0.852549696	1.051454666	1.299490326	1.910960195
J_{44}	0.597444858	0.842776110	1.030489463	1.275986504	1.913979709
J_{45}	0.601222195	0.844669241	1.028788249	1.275986504	1.913979709
J_{46}	0.595460368	0.845894251	1.033453225	1.261745563	1.913979709
J_{47}	0.588814584	0.842776110	1.027148308	1.265043937	1.910960195
J_{48}	0.590662450	0.850657669	1.035278366	1.275986504	1.921742678
J_{49}	0.598372424	0.862337174	1.043399683	1.262927700	1.934141825
J_{50}	0.591030413	0.851508322	1.035278366	1.257473805	1.974699741
J_{51}	0.592397659	0.842936838	1.035278366	1.275231717	1.986756456
J_{52}	0.575729560	0.854603003	1.034199823	1.284230276	1.958125531
J_{53}	0.574158805	0.840822441	1.031918443	1.264579325	1.971260512
J_{54}	0.571549022	0.837721413	1.018490066	1.275348100	1.955603380
J_{55}	0.565836095	0.838220617	1.033453225	1.278164842	1.987566444
J_{56}	0.560519145	0.834560214	1.020789203	1.263601985	2.009223412
J_{57}	0.570354465	0.818330276	0.993699316	1.259025884	1.919270208
J_{58}	0.566708552	0.827973127	0.995205761	1.232760613	1.976734648
J_{59}	0.562409792	0.835161909	0.991687686	1.213715589	1.943266795
J_{60}	0.573288049	0.824682942	0.992062779	1.222188079	2.009503467
J_{61}	0.567814006	0.857883160	1.031918443	1.254869129	2.060114657
J_{62}	0.575595834	0.849653188	1.050129374	1.262336586	2.084111243
J_{63}	0.609652318	0.891687833	1.095787160	1.333614431	2.155842844
J_{64}	0.611094040	0.807900162	0.993699316	1.217998479	2.135917358

Notes: Parameters directly estimated from GSOEP.

7.4.3 Calibrated Parameters

Table 12: Data Targets

Param.	Description	Data Targets
<i>Health Transition and Acute Illness</i>		
Π	Prob. to drop in health status after severe illness cond. on health state	Health state at end of working life for top income group
S^{sev}	State of sickness that threatens health	Health state at end of working life for top income group
<i>Sickness Aggravation</i>		
ζ	Prob. untreated sickness becomes severe	Extensive Margin for Bottom Income Quintile
κ	Aggravation Factor	Intensive Margin for Bottom Income Quintile
<i>Preferences</i>		
ψ	Multiplier of marginal utility	Top income quintile worker do not go sick to work
<i>Government</i>		
τ	Proportional Tax Rate	Balanced Government Budget

Notes: Calibrated parameters and their data targets.

Table 13: Calibrated Parameters

Param.	Value	Statistics	Data	Model
<i>Health Transition</i>				
Π_1	1.0	Prob. to be in H_1 for $K_5 J_{55}$	1.0%	1.8%
Π_2	0.9825	Prob. to be in H_2 for $K_5 J_{55}$	10.7%	10.1%
Π_3	0.965	Prob. to be in H_3 for $K_5 J_{55}$	35.8%	35.9%
Π_4	0.89	Prob. to be in H_4 for $K_5 J_{55}$	45.2%	44.8%
Π_5	0.645	Prob. to be in H_5 for $K_5 J_{55}$	7.3%	7.4%
S^{sev}	10			
<i>Sickness Aggravation</i>				
ζ	0.25			
κ	0.5			
<i>Preferences</i>				
ψ	0.001			
<i>Government</i>				
τ	0.21	Balanced Budget		

Notes: Calibrated Parameter values and match of model to data.