

Lorenz, Olga; Goerke, Laszlo

Conference Paper

Commuting and Sickness Absence

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Labor - Empirical Studies 5, No. G16-V4

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Lorenz, Olga; Goerke, Laszlo (2015) : Commuting and Sickness Absence, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Labor - Empirical Studies 5, No. G16-V4

This Version is available at:

<http://hdl.handle.net/10419/113173>

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.

Commuting and Sickness Absence^{*}

Laszlo Goerke

IAAEU - Trier University ⁺, IZA, Bonn and CESifo, München

⁺ Institute for Labour Law and Industrial Relations in the European Union, Campus II
D – 54286 Trier, Germany
E-mail: goerke(at)iaaeu.de

Olga Lorenz

IAAEU - Trier University ⁺⁺

⁺⁺ Institute for Labour Law and Industrial Relations in the European Union, Campus II
D – 54286 Trier, Germany
E-mail: lorenz(at)iaaeu.de

Preliminary draft, August 25, 2015

Abstract: In this paper, we focus on the causal effect of commuting on sickness absence from work using German panel data. According to theory, the effect of commuting on the number of workers' absence days may be positive or negative. Empirical tests of this effect are not standard, due to reverse causation and lack of good control variables. To address reverse causation, estimates of commuting on sickness absence are derived using changes in commuting distance for employees who stay with the same employer and who have the same residence during the period of observation. Our results show that employees, who commute long distances, are absent more often than employees with shorter commutes. We explore explanations for the effects of long distance commutes to work and can find no evidence that it is due to working hours mismatch, reduced leisure time or differences in health status.

Keywords: sickness absence, absenteeism, commuting, health, labour supply

JEL: I10, J22, R2, R40

^{*} We thank Mario Mechtel and the participants of the 7th Economic Workshop in Trier, the 2015 Colloquium on Personnel Economics (COPE) in Vienna, the 2015 Spring Meeting of Young Economists (SMYE) in Ghent, the 2015 annual conference of the Verein für Socialpolitik (VfS) in Münster and seminars at U Trier for helpful comments and suggestions.

1. Introduction

Each day millions of employees commute between home and work. Moreover, the frequency of commuting and the average duration of commutes have risen in the last decades. Hence, commuting is an important and growing component of daily life. According to the German Federal Statistical Office, the percentage of commutes which took less than 10 minutes to work is declining, while the share of those people who commute 30 to 60 minutes to work has risen from 17% in 1996 to 23% in 2012. A similar picture emerges regarding commuting distance (German Federal Statistical Office, 2013). This trend is not unique to Germany. In the UK commuting times have increased from 48 to 54 minutes per day, the average commuting trip length has increased by 1.3 kilometres between 1995/1997 and 2012 to reach 14.5 kilometres (National Travel Survey 2012). In Spain and Italy, over the period 1997 – 2006 commuting times have increased from 31 to 34 minutes and 22 to 35 minutes, respectively.¹

On the one hand, commuting may be viewed positively as it increases the density of labour markets and, hence, allows for better matches between jobs and individuals. Moreover, commuting enables employees to live in places where there are no adequate jobs, without forsaking their income. On the other hand, commuting is usually argued to be problematic from an environmental point of view and to be detrimental to the health of employees.²

If commuting is negatively related to health, there may also be a relationship between commuting and absence from work. If commuting makes people ill, employees who commute are more likely to be absent from work (Zenou, 2002). In addition, the gain from absence in terms of hours which can be used for other purposes than work, such as recuperation, are likely

¹ Data is from the European Survey on Working Conditions (EWCS).

² Several predominantly U.S. studies have found that work commutes induce stress due to its unpredictability and the perceived loss of control (Gottholmseder et al., 2009). Furthermore, commuting has been shown to be associated with increased heart rate and blood pressure (Novaco et al., 1979; Schaefer et al., 1988) as well as a higher body mass index and a higher percentage of body fat (Flint et al., 2014). Further, commuting translates into shorter sleeping times and sleep disorders (Costa et al., 1988; Walsleben et al., 1999; Hansson et al., 2011), a lower social capital and participation (Mattisson, 2014), which has in turn been associated with health outcomes (Putnam, 2000; Lindström, 2004; Besser et al., 2008), negative mood (Gulian et al., 1989), emotional arousal (Hennessy and Wiesenthal, 1997), lower general well-being and life satisfaction (Stutzer and Frey, 2008; Olsson et al., 2013) as well as higher levels of workplace aggression (Hennessy, 2008), poor concentration levels (Matthews et al., 1991) and a higher risk of mortality (Sandow et al., 2014).

to be higher for individuals who commute. However, from a theoretical vantage point, the effect of commuting on sickness absence may also be negative. Individuals would not choose to have a longer commute unless they were compensated for it, either in the form of improved job characteristics (including pay) or better housing prospects (Stutzer and Frey, 2008). Hence, individuals who commute may have better, more motivating jobs and be able to achieve a better work-life balance. Furthermore, willingness to travel to work may be associated positively with work effort. Accordingly, the net effect of the commute to work on sickness absence is an empirical issue.

From an economic point of view, an understanding of the impact of commuting on absence behaviour and an analysis of the determinants of absence from work is important for a number of reasons: First, if absence affects productivity and profitability, a firm's employment but also location decision may be influenced by its (prospective) employees' commuting behaviour. Second, since absence causes an externality, at least if absent employees are entitled to sick pay, health policy requires knowledge of the relationship between commuting and absence from work. Third, policies which alter mobility of workforce and the integration of economic regions needs to take into account the effects of commuting on absence. Finally, this analysis broadens our perspective on the economic costs of commuting beyond the direct monetary costs such as infrastructure and commuting expenses.

Although, therefore, knowledge of the impact of commuting on absence behaviour is valuable for various reasons, there is only limited empirical evidence. Moreover, many of the extant studies examine correlations without considering intervening or mediating factors, are based on cross-sectional data from a single workplace as well as small and relatively restricted samples, and do not investigate whether the determinants of absence differ between commuters and non-commuters. Against this background, this paper contributes to the literature in four ways: First, using German panel data (SOEP) for the years 2002 – 2011 we examine the causal effect of the length of the commute on the number of absence days by analysing the impact of

employer-induced changes in commuting distance. In particular, we look at workers who stay with the same employer and have the same place of residence during the period of observation. Keeping the workers' employer and residence constant allows us to address endogeneity of commuting distance. Thus, this paper improves upon preceding work by using a more sophisticated statistical approach than in most of the previous related contributions, and consequently, provides more reliable results and allows isolating the causal effect of the work commute. Second, we use, *inter alia*, a fixed effects framework that includes variables known to determine absence from work as well as factors which may provide compensation for commuting, such as income, job satisfaction and housing quality. In this way, we deal with time-invariant unobserved worker characteristics and reduce reverse causation. Moreover, by controlling for all important determinants of sickness absence and by considering factors which may provide compensation for commuting in a single regression, we present a more integrated approach for explaining the relationship between sickness absence and commuting than found in the previous literature. Third, we are able to enquire whether absence behaviour of employees who do not commute differs from that of employees who choose to make short, middle or long distance commutes to work. As we use a sample including individuals working for a large variety of employers, our results are more generally applicable than those of other studies, most of which have concentrated on single firms. Fourth, this paper investigates some potential channels driving the relationship between commuting and sickness absence to obtain a fuller picture of how commuting affects behavioural (lifestyle) factors that affect absence behaviour.

Our empirical analyses show that, even after all the aspects mentioned above are considered, commuting still has an important effect on sickness-related absence from work. We find that employees, who commute long distances (more than 50 kilometres), are absent more often than comparable employees who do not commute or who travel shorter distances. In particular, the average number of absence days amounts to 10.4 days for the entire sample, while long distance commuters exhibit 11.9 absence days. These descriptive findings are confirmed

when accounting for observable characteristics in a pooled sample as well as in the panel structure of our data. Our results suggest that especially long distance commutes significantly increase the absence days of employees, and this result is robust to numerous different specifications.³ We explore possible explanations for the effect of long distance commutes on absence from work and can find no evidence that it is due to working hours mismatch, reduced leisure time or differences in health status.

The remainder of this paper is organized as follows: Section 2 considers the literature on commuting and sickness absence. Section 3 describes the data and variables. Section 4 focuses on our identification strategy and sets the econometric method. Section 5 reports the results, including several robustness checks and the analysis of mechanisms through which commuting might affect individual's absence behaviour. Finally, section 6 concludes the study.

2. Background

Commuting has attracted substantial attention in economics, at least since Becker (1965) considered it in his model of optimal time allocation. In transport economics the focus is on choice of transportation modes, in urban economics the emphasis lies on residence location. In labour economics, commuting is seen as a source of labour mobility and has been incorporated into models of job search (van den Berg and Gorter, 1997; van Ommeren, 1998; Rouwendal, 2004) and labour supply (Wales, 1978; Gutiérrez-i-Puigarnau and van Ommeren, 2010; Gershenson, 2013; Gimenez-Nadal and Molina, 2014). This paper complements the literature by focusing on the causal effect of commuting on sickness absence from work, which is an important aspect of labour supply.

Absenteeism or sickness absence has been the focus of much research by social scientists and economists (see Brown and Sessions, 1996; Treble and Barmby, 2011 for surveys). To date, few studies have looked at the effects of commuting on sickness absence. Knox (1961)

³ Note that in this paper we use *absenteeism* and *sickness absence* as synonyms since it is quite impossible to distinguish between voluntary and involuntary absence.

establishes a positive correlation between being absent frequently and living more than 60 minutes away from work for 300 male employees in a large cotton factory near Buenos Aires. Taylor and Pocock (1972) investigate the relationship between commuting time and absenteeism in a population of 1,994 office workers in Central London. The results indicate that workers who travelled more than one hour and a half to work showed a greater absence frequency than those who had a shorter duration of travel. In an analysis of 79 employees in California Novaco et al. (1990) find evidence of a positive correlation between commuting and absence which, however, was not robust to the inclusion of control variables. Magee et al. (2011) employ data from the Australian household, income, and labour dynamics data set (HILDA) for the years 2005 and 2008. They find a positive correlation between commuting time and absence, which is particularly strong for those whose total commuting time exceeds five hours per week. Finally, Künn-Nelen (2015) uses the British Household Panel Survey (BHPS) data for 1991 to 2008 and finds no robust correlation between commuting time and being absent during the week prior to the interview.⁴

In sum, the literature points to a relationship between commuting and sickness absence. However, many of the relevant studies only examine correlations, without sufficient consideration of intervening or mediating factors. Furthermore, most analyses are based on cross-sectional data from a single workplace and/or small and relatively restricted samples. Consequently, it is not clear whether the results can be generalized (cf. Dionne and Dostie, 2007). The study our analysis is most closely related to is the one by van Ommeren and Gutiérrez-i-Puigarnau (2011). They examine the impact of commuting distance on workers' productivity using the German Socio-Economic Panel (SOEP). The authors argue that the number of days absent from work is closely related to shirking behaviour and is, therefore, a reasonable and inverse measure of workers' productivity. van Ommeren and Gutiérrez-i-Puigarnau (2011) adopt an extended urban efficiency wage perspective, assuming that a worker's

⁴ Moreover, there are some empirical analyses of absence behaviour which include an indicator of commuting as covariate, without looking at the relationship in detail. Allen (1981) and De Paola (2010), for example, report no correlation.

shirking decision depends, among other things, on residence location, which is chosen by an individual, and thus, on commuting distance. The authors address the resulting endogeneity issue by using residence fixed effects and changes in commuting distances that are employer-induced.⁵ They find that commuting distance induces shirking behaviour with an elasticity of about 0.07 to 0.09. While substantially enhancing our knowledge of the impact of commuting on sickness absence, the study still has several limitations. First, van Ommeren and Gutiérrez-i-Puigarnau (2011) have omitted important confounding factors, such as gender, marital status, educational attainment, job tenure and job satisfaction.⁶ Failure to include these confounders into the estimations is likely to result in overestimates of the strength of the association between commuting and sickness absence. Second, rational individuals will only choose to commute if they are compensated.⁷ Compensation may be provided in the housing or labour markets. Therefore, in what follows, we also consider how job satisfaction, satisfaction with leisure time and housing quality (incorporating satisfaction with dwelling and a household crowding index) affect the employees' sickness absence behaviour along-side commuting distance. Third, van Ommeren and Gutiérrez-i-Puigarnau (2011) restrict their sample only to individuals who commute. Obviously, these sample restriction is a source of a potential selection bias. Fourth, since van Ommeren and Gutiérrez-i-Puigarnau (2011) ignore observations with zero commutes, they are not able to investigate whether the determinants of absence differ between commuter and non-commuter. Fifth, the authors estimate a log-linear relationship between commuting distance and absence and report that the results are “hardly sensitive” (p. 3) to this assumption. This specification does not allow for discontinuities in the impact of commuting on absence behaviour, as we observe it in our data. Finally, given the efficiency wage perspective, van

⁵ This identification strategy has been used by other authors, as well, looking at different issues. Zax (1991) and Zax and Kain (1996) analyse job and residential moving behaviour. Gutiérrez-i-Puigarnau and van Ommeren (2010) investigate labour supply patterns and Roberts et al. (2011) consider the effect of an exogenous change in commuting time on psychological health in a robustness check. Finally, Carta and De Philippis (2015) investigate the impact of commuting on the labour supply of couples.

⁶ See e.g. Dionne and Dostie (2007) who provide new evidence on the determinants of sickness absence.

⁷ According to standard economics, commuting confers disutility and therefore, the burden of commuting is only chosen when traveling longer distances to and from work is either compensated by an intrinsically or financially rewarding job or by additional welfare gained from a pleasant living environment (see e.g. Stutzer and Frey, 2008).

Ommeren and Gutiérrez-i-Puigarnau (2011) interpret absence as a proxy for shirking and are not concerned with validating this or other explanations for the observed effect. This paper will proceed, inter alia, by testing potential channels driving the relationship between commuting and sickness absence.

3. Data and Variables

The current study is based on information from the German Socio-Economic Panel (SOEP). The SOEP is a longitudinal, nationally representative survey of private households in Germany. Currently, around 30,000 people in approximately 15,000 households participate in the survey. The SOEP includes rich information on labour market status, wealth, incomes and standard of living; health and life satisfaction as well as on family life and socio-economic variables.⁸ To the best of our knowledge, the SOEP is the only person-level dataset for Germany providing detailed information on both absence from work and commuting distance.

The information with respect to sickness absence is derived from the SOEP “Health and Illness Questionnaire” (HIQ). The HIQ provides a self-reported measure of the annual number of days absent from work due to sickness in the previous year. The exact question reads as follows: *“How many days were you unable to work in 20XX due to illness? Please state the total number of days, not just the number of days for which you had an official note from your doctor: (a) None (b) A total of X days.”* The great advantage of this question is that it provides information on the *total* number of absence days, and not only with respect to those, for example, for which a medical certificate is required.⁹ However, there is no data in the SOEP on the annual number and length of specific sickness spells. Given the SOEP questionnaire, our dependent variable of

⁸ Further information about the SOEP is provided by Wagner et al. (2007) and can also be found at: <http://www.diw.de/english/soep/29012.html>. We use the SOEP long v29 dataset.

⁹ In Germany, dependent employees with a minimum tenure of four weeks can basically take sick leave without a durational restriction. From the third day of the sickness spell onwards, employees are legally required to present a confirmation by a doctor certifying an illness which prevents the employee from attending work, but employees are obliged to inform their employers about their sick leave from the first day of their absence. During the first six weeks of an absence period, the employer has to continue to pay wages. These continued wage payments are subject to income taxation. Once an employee’s absence period exceeds six weeks, the mandatory health insurance will cover the cost of sick pay which drops to at most 90% of the net wage.

interest is labelled ‘days absent’ and counts the total number of days the employee has been absent during the previous year.

The SOEP, furthermore, requires respondents to report on commuting distance. The question reads: “*How far (in kilometres) is it from where you live to where you work? (a) X km (b) Difficult to say, location of workplace varies (c) Workplace and home are in the same building/same property.*” We define all respondents for whom either part (c) of the question applies or who state that the distance between home and workplace is less than ten kilometres (part (a) of the question) as non-commuters. All respondents who travel ten or more kilometres to work (part (a) of the question) are defined as commuters. In particular, those who travel to work between ten kilometres and less than 25 kilometres are short distance, those who travel 25 kilometres and less than 50 kilometres are middle distance and those who cover 50 or more kilometres are long distance commuters.¹⁰ This approach allows for qualitatively different effects of, for example, shorter and longer commuting distances on absence. Moreover, it is not sensitive to minor reporting errors.¹¹ Finally, those who report working in different places (part (b) of the question) were excluded from the analysis as it is difficult to determine their actual commuting distance.

It is worth mentioning that the SOEP provides direct information about commuting time and commuting mode only in 2003. In addition, it is possible to imprecisely ascertain commuting time indirectly by calculating the difference between daily working hours including travel time to and from work and the usual daily working hours. We use this information in a robustness check (cf. Section 5.4). However, in our context, commuting distance is more appropriate because commuting time may be influenced by many reasons like an increase in congestion, transportation mode or even by an improvement in public transportation infrastructure.

¹⁰ No standard definition of commuting is used internationally or in Germany. We build our categories in line with definition used by the Federal Statistical Office of Germany.

¹¹ Furthermore, we have experimented with several functional forms and categorisations for distance, for instance, we employed commuting distance as continuous or as dichotomous variable. The main results are hardly sensitive to the chosen specification.

The choice of the other explanatory variables is informed by the literature on the determinants of sickness absence as well as on commuting. Correlates or determinants of absence can be categorized as follows (e.g. Block et al. 2014; Livanos and Zangelides, 2013; Ziebarth and Karlsson, 2010; Frick and Malo, 2008; Dionne and Dostie, 2007). The first group contains variables on personal characteristics such as gender, marital status, children, age and age squared, current health status as well as educational attainment. The second set incorporates variables on job-related aspects: Tenure, working time, type of employment contract (temporary), occupational position, size of company, sector information, industry dummies, and income. We also include state as well as year dummies. Furthermore, studies on commuting (e.g. Costa et al., 1988; Stutzer and Frey, 2008; Lyons and Chatterjee, 2008) suggest that compensation for commuting may be provided in the housing or labour market. Hence, we also include indicators of satisfaction with dwelling and the amount of leisure time, job satisfaction and a household crowding index as explanatory variables.

We restrict our sample to the years 2002 – 2011 since previous waves of the SOEP which contain a question on commuting distance do not provide this information for respondents who reside and work in the same town. Our estimation sample consists of 18 to 65-year-old individuals in paid employment and does not include self-employed.¹² As information on sickness absence refers to the year before the interview date ($t - 1$) and commuting distance is measured at the interview date (t), we have to ensure that the commuting distance reported at the interview date applies to the same year for which sickness absence days were reported. Accordingly, to ensure that the given information applies to the same year, we first matched the retrospective information which we are interested in with the current information of the relevant year. Furthermore, to affirm that information on commuting distance and sickness absence refers to the same employer we additionally confine our sample to workers who have been working for

¹² See, for example, Roberts et al. (2011) for a comparable approach.

at least two years for the same employer.¹³ As part of our identification strategy explained in the next section we want to assure that the effect of commuting on absence behaviour is not due to a change of residence or employer. Hence, we finally restrict our sample to respondents who stay with the same employer and who have the same residence during the whole period of observation. This results in a working sample of 6,459 individuals with 31,567 observations.

Table A.1 in the Appendix shows our variable definitions. For a complete list of covariates and descriptive statistics see Appendix A Table A.2.

4. Identification Strategy and Econometric Methods

4.1 Identification Strategy

Our aim is to test the causal effect of commuting on absence behaviour of workers. Usually, the worker's commuting distance is self-chosen by the worker and thus, determined by the endogenously chosen residence location and the workplace location. Hence, it is difficult to interpret the effect of commuting distance on sickness absence as a causal effect. A consistent estimation requires that commuting distance, more precisely the change of commuting distance is exogenous. Therefore, we select workers who did not change employer or residence during the period of observation. Hence, we keep residence location and employer constant. It is plausible that if a worker does not change employer and does not change residence while a change in commuting distance is observed in the data, then the worker must have change workplace location because e.g. of a firm relocation. So, changes in distances are due to exogenous shocks.¹⁴ Such changes in workplace location due to firm relocation have been shown to be quite common (Gutiérrez-i-Puigarnau and van Ommeren, 2010; Gutiérrez-i-Puigarnau and van

¹³ Since we use workers' fixed effects, it is unlikely to create a selection bias because the FE specification controls for worker-specific time-invariant heterogeneity.

¹⁴ It is important to note, that in the data available there is no information on whether the worker's firm relocated or not. So it is not possible to distinguish between true changes (because of firm relocation) and misreporting. Since we are treating commuting distance as a categorical variable our results are not sensitive to minor reporting errors. Hence, the downward bias in our estimate is likely to be small. We additionally address this problem by excluding observations referring to absolute changes in commuting distance smaller than two kilometres (cf. Section 5.4). The results are almost identical.

Ommeren, 2013). For example, about 16.5% of firms in Germany are each year involved in relocation decisions (Federal Statistical Office of Germany, 2008). Using this approach, in our sample about 10% of changes in commuting distance are employer-driven. Selecting a sample of workers who do not change employer and who do not change residence may create a selection bias. We will explicitly address the potential bias of this selection by comparing results of different samples (cf. Section 5.4).

4.2 *Models for Cross-Sectional Data*

Since absence days can only take on non-negative integer values we estimate a negative binomial model with heteroskedasticity robust standard errors, which is also a convenient way for dealing with overdispersed data, such as we are examining.¹⁵ Additionally, we make use of an OLS regression model with heteroskedasticity robust standard errors. This approach is feasible since we need not get the functional form perfectly right to obtain valid estimates of the average partial effects. The idea for the empirical test is captured in the following regression equation (1):

$$A_i = \alpha + \beta D_i + \gamma X_i + \varepsilon_i \tag{1}$$

Here, A_i is a measure of the total number of days absent at work for individual i . D_i is an indicator for commuting distance and X_i represents the different vectors of independent variables (e.g. variables on personal characteristics, variables on job-related aspects and variables compensating for commuting). In order to capture state and time specific effects we also include state and year dummies. β and γ are coefficients, and ε_i denotes the error term. Our main interest lies in β . The pooled estimators identify the effect of commuting on the reported number of days absent, based on the variation in these variables between people and for each individual over time. It is assumed that unobserved characteristics, as well as measurement errors, are captured in the error term of the estimation.

¹⁵ The overdispersion parameter corresponds to $\alpha = 2.73$.

4.3 Models for Panel Data

Additionally, we assess the impact of change in commuting distance on a change in the outcome variable using fixed effects models because causal inference is better supported using panel data, rather than cross-sectional data (Wooldridge, 2010). Accordingly, we can eliminate the risk that time-invariant variables may confound the relationship between commuting distance and sickness absence.

Since our dependent variable is a count variable, we employ a conditional fixed-effects negative binomial regression model as a benchmark.¹⁶ Conditional estimation of the fixed-effects model is obtained using maximisation of the log likelihood conditional on the sum of the number of counts during the period during which the individual is observed. Although the conditional fixed-effects negative binomial method is frequently used, it has been criticised as not being a true fixed-effects model since it fails to control for all of its stable predictors (Allison, 2009; Allison and Waterman, 2002). The unconditional negative binomial model and the multinomial model have been suggested as alternatives. Since the negative multinomial model fails to handle overdispersion and the unconditional model is unsuitable for large data sets with lots of variables, we revert additionally to the fixed-effects OLS regression.¹⁷ Furthermore, a fixed-effects OLS regression is less contingent on distributional assumptions and easier to interpret than the alternatives. The basic model specification can be denoted (2):

$$A_{it} = \beta D_{it} + \gamma X_{it} + \alpha_i + \alpha_t + \varepsilon_{it} \quad (2)$$

where A_{it} is a measure of the number of days absent for a worker i in a specific residence and with a specific employer in a year t , D_{it} is an indicator for commuting distance, X_{it} are a set of conditioning variables, β and γ refer to parameters to be estimated. α_t are defined as year fixed effects and α_i are individual fixed effects for each residence and employer combination.

¹⁶ For detailed explanation, see Hilbe (2007).

¹⁷ The unconditional negative binomial model leads to underestimated standard errors, which yields biased estimates.

5. Results

5.1 Sample Description

Table 1 reports, inter alia, the associations between commuting distance, the number of days absent and the incidence of absence in the data. The table shows that in our sample the average number of days lost through sickness absence amounts to 10.36 days. The standard deviation of 24.7 days indicates that there is a lot of cross-sectional variation. The distribution of absence days is heavily skewed right with a mass point at zero. The full distribution can be found in Appendix A Figure A.1.

Table 1

Descriptive statistics for full sample and for commuter categories. Source: SOEP 2002 – 2011.

	Full sample				Non-Commuter		Short distance commuter		Middle distance commuter		Long distance commuter	
	Mean	SD	Min	Max	Mean	SD	Mean	SD	Mean	SD	Mean	SD
<i>Absence:</i>												
# of days	10.36	24.70	0	365	9.96	24.31	10.59	25.12	10.43	23.72	11.86	27.62
Incidence	0.65	0.47	0	1	0.63	0.48	0.65	0.47	0.67	0.46	0.70	0.45
<i>Focal variable:</i>												
NC	0.46	0.49	0	1								
SDC	0.33	0.47	0	1								
MDC	0.16	0.36	0	1								
LDC	0.05	0.23	0	1								
N	31,567				14,113		10,435		5,129		1,890	

Notes: Summary statistics only for key variables. NC = Non-commuter, SDC = Short distance commuter, MDC = Middle distance commuter, LDC = Long distance commuter, SD = Standard deviation. Appendix A shows the detailed descriptive statistics in Table A.2.

About one-half (54%) of individuals in our dataset are short, middle or long distance commuters. Sickness absence days vary widely within the group of workers, who chose to commute to work. The annual number of days absent increases by about two days, as one-way commute distance increases from under 10 kilometres (non-commuter) to over 50 kilometres (long distance commuter). Those workers who commute long distances have on average 11.9 absence days. The incidence rate is also higher. Approximately 70% of long distance commuters have stayed home sick at least once in the last 12 months, whereas 63% of non-commuters did so. Hence, the descriptive evidence suggests that, being a commuter is associated with a higher

incidence of absence and more absence days per year.¹⁸ The descriptive statistics furthermore indicate that commuters are more often male, are better educated, work longer hours, and are less likely to work part time. Furthermore, commuters have a higher labour income, a shorter tenure and tend to work more often in large firms with more than 200 employees. Finally, commuters tend to be less satisfied with their leisure time and with their work than non-commuters.¹⁹

In our data, the average one-way commuting distance of workers is 19 kilometres, so the daily commuting distance travelled is 38 kilometres. This is in line with a range of other studies employing German data (OECD, 2007; Schulze, 2009; Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR), 2012). Hence, our sample selection is likely unrelated to commuting behaviour. The full distribution can be found in Appendix A Figure A.2.

5.2 Cross-sectional Evidence

Results from various multivariate regression models with the number of days absent as the dependent variable and a categorical variable for commuting distance as focal variable are reported in Table 2.²⁰

Table 2

Estimation results. Dependent variable: Days absent. Source: SOEP 2002 – 2011.

	Model I Pooled NEGBIN	Model II Pooled OLS	Model III FE NEGBIN	Model IV FE OLS
Short distance commuter	0.0385 (1.45)	0.572 (1.86)	0.0442 (1.91)	1.305 (1.94)
Middle distance commuter	0.0705* (2.17)	0.846* (2.17)	0.109*** (3.65)	2.607** (2.79)
Long distance commuter	0.201*** (3.72)	2.173*** (3.36)	0.191*** (4.35)	3.370* (2.56)
<i>N</i>	31,567	31,567	31,567	31,567

Notes: Only the coefficients for the commuting variables are reported. Non-commuters are treated as the reference category. The following control variables are included: female, age, age squared, married, children, college degree, education, health status, working hours, regular part-time, temporary job, blue-collar worker, firm size, public service, tenure, tenure squared, log(monthly wage), satisfaction with work, satisfaction with leisure, satisfaction with dwelling, HCI, business sector dummies, region dummies, year dummies. Appendix A shows the results for control variables in Table A.3. *t* statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

¹⁸ In the following multivariate regression we only consider the number of workdays lost due to illness, which is judged to be economically more relevant, compared to the incidence of absence.

¹⁹ Appendix A shows the detailed descriptive statistics in Table A.2.

²⁰ Appendix A shows the results for control variables in Table A.3.

Model I estimates a pooled negative binomial regression (NEGBIN) with robust standard errors, taking a large number of individual characteristics and job related aspects into account, as well as year and region dummies. Model I shows that sickness absence is higher the longer the commute, *ceteris paribus*. While the estimated coefficient of being a short distance commuter is insignificant, we find that individuals who commute middle or long distances to work have a longer duration of absence compared to non-commuters.²¹ Being a middle distance commuter instead of a non-commuter is associated with a 7.05% change in the expected number of days absent, or equivalently, the conditional mean is 1.07 times larger. Being a long distance commuter leads to a 0.201 proportionate change or 20% change in the number of sickness absence days. The effect is, for example, greater than the impact due to being a female. Model II estimates a pooled least squares regression (OLS). The regression results are almost identical to the ones reported above, indicating that greater commuting distances are associated with more sickness absence. For example, long distance commuters are on average about 2.17 days more absent than those who commute less than 10 kilometres. Since the raw difference in the duration of absence observed between long distance commuters and non-commuters is 1.90 days (Table 1), this difference tends to underestimate the impact of commuting.

5.3 *Fixed-effects Analyses*

Since non-commuters, short distance, middle distance and long distance commuters may differ in time-invariant, unobservable characteristics which affect absence behaviour, we next present the findings from fixed effects estimations.

In Table 2, Model III reports the results for a fixed-effects negative binomial estimation (FE NEGBIN). Using the conditional method, Model III shows that the overall effect of

²¹ Recall that the dependent variable is a count variable, and negative binomial regression models the log of the expected count as a function of the predictor variables. We can interpret the negative binomial regression coefficients as follows: for a one unit change in the predictor variable, the difference in the logs of expected counts is expected to change by the respective regression coefficient, given the other predictor variables in the model are held constant. Hence, the coefficients displayed in Model I and Model III (Table 2) are equal to the proportionate change in the conditional mean if the regressors change by one unit. For indicator variables the coefficient reflects a proportionate change from the base level. For detailed explanation, see Cameron and Trivedi (2008).

commuting distance on the number of days absent is positive and statistically significant at the 0.01% level.²² Comparing the magnitudes of the estimated coefficients in Models I and III clarifies that controlling for time-invariant characteristics does not take away but rather increases the effect of being a middle or long distance commuter on the number of days absent. Hence, cross-section estimation of the effect of commuting on sickness absence negatively biases the results. One plausible explanation for this bias is that individuals with unobserved positive attitudes to work are more likely to accept jobs at longer distances and are also less likely to be absent. So, the conditional method estimates reported here are conservative. If we consider middle and long distance commuters, the most highly statistically significant regressors, the coefficients indicate that the expected number of days absent is about 11% higher for middle distance commuters and 20% higher for long distance commuters compared to non-commuters.

Since the conditional negative binomial method has been criticised for not being a true fixed effects model, we also estimate a fixed-effects least square model (FE OLS) with robust standard errors (Model IV). The coefficients of interest are identified by those individuals who change their commuter status at least once from being a non-commuter to being a short distance, middle distance or long distance commuter, or vice versa. Model IV in Table 2 shows that the overall effect of commuting distance is positive and statistically significant.²³ Again, comparing the magnitudes of the estimated coefficients in Models II and IV shows that, even after controlling for individual fixed-effects, the effect of commuting on sickness absence increases. Being a long distance commuter instead of a non-commuter is associated with 3 absence days more on average ($p < 0.05$), while being a middle distance commuter goes along with 2 more absence days on average ($p < 0.01$).

²² The three degree-of-freedom chi-square test indicates that commuting distance is a statistically significant predictor of absence ($\chi^2(3) = 25.27; p = 0.0000$).

²³ The F-test indicates that commuting distances are jointly significant at the 5% level ($F(3, 6458) = 3.38; p = 0.0175$).

Thus, the descriptive evidence and the results of the pooled estimations are confirmed: While short distance commutes have no impact on sickness absence, middle and long distance commutes increase the duration of absence.²⁴

Running separate fixed effects negative binomial regressions for long, middle, short distance commuter and non-commuter enables us to have a closer look at the coefficients of the explanatory variables and examine whether these variables play different roles for the four different groups (not reported). Starting with the group of variables on personal characteristics it is apparent that the variables reflecting gender, age, being married or having a partner, children living in the household, educational attainment and the current health status seem to play a similar role for the absence of long, middle, short distance commuter and non-commuter. The respective variables are jointly statistically significant at the 0.01% level for commuters as well as for non-commuters. Looking at the college degree, for long and middle distance commuter having a completed college education is associated with a reduction of sickness absence days, *ceteris paribus*. Since the 95% confidence intervals for both groups overlap, they do not seem to be statistically significantly different. In contrast, this effect is bigger and statistically significantly different for short distance commuter, indicating that having a college degree results in more absence days compared to long and middle distance commuter. Regarding the variable reflecting the current health status, it is recognisable that a better state of health is associated with less absence days for all four groups. The dummies for self-reported health status are jointly statistically significant at the 0.01% level for (long, middle, short distance) commuters as well as for non-commuters. Long distance commuters seem to react less sensitive to health status, but the 95% confidence intervals for all groups overlap, suggesting that they are not statistically significantly different.

²⁴ We also estimated a random-effects probit model where we only distinguish between *never having been absent* and *having been absent at least once*. The regression shows a positive and significant effect. Being a middle or long distance commuter increases the probability of being absent ($\beta_{long\ distance} = 0.1407$ with $p = 0.005$ and $\beta_{middle\ distance} = 0.0779$ with $p = 0.020$). This also supports the hypothesis that only longer commuting distances positively affect absence from work.

In contrast to variables on personal characteristics, the group of variables on job-related aspects do not seem to play a similar role for the absence of commuters and non-commuters. For long and middle distance commuters, the respective variables (tenure, working hours, part-time work, temporary job, firm size, etc.) are not jointly statistically significant. In particular, none of the job-related variables mentioned above seems to be of importance for the absence of work of long distance commuters. The same variables are statistically significant at the 0.01% level for short distance commuters and non-commuters. For both groups working hours, firm size, income and working in public sector seem to be positively related to absence from work. Non-commuters seem to react more sensitive to firm size and income (e.g. reporting to work in a firm with more than 2,000 employees is associated with 52% more absence days for non-commuters and only 37% more absence days for short distance commuter), but the 95% confidence intervals for both groups overlap, suggesting that they are not statistically different.

Finally, variables which may provide compensation for commuting (satisfaction with dwelling, satisfaction with the amount of leisure time, job satisfaction, and household crowding index) do not seem to play a similar role for absence of the four groups. For long distance commuter and middle distance commuter we do not find a jointly statistically significant effect of the respective variables, whereas we do find a jointly significant effect for short distance commuter and non-commuter, suggesting that e.g. satisfaction with work, satisfaction with leisure or satisfaction with dwelling provide at least compensation for the commute to work up to a certain limit of commuting distance.

All in all, the analysis shows that personal characteristics, job related aspects and variables which may provide compensation for commuting are important factors associated with sickness absence from work of both commuters and non-commuters. Still, differences in these variables between these groups do not contribute much to explain the difference in absence between long, middle, short distance commuter and non-commuter.

5.4 *Robustness Checks*

In Table A.4 (Appendix A) we report a number of estimates designed to act as robustness checks. We focus on the variables of interest for the fixed-effects negative binomial model, which is depicting the estimated coefficients for the three commuting categories.

The first two models ((i) and (ii)) are estimated for men and women separately as the effect of commuting on absence may be gender-specific (Leigh, 1993; Vistnes, 1997). The estimated coefficients for men and women are broadly the same and not statistically different from each other. Hence, we obtain no evidence that the effect of commuting distance on absence is gender specific.

The third model (iii) is estimated for those individuals who do not work in the public sector as it is known that the number of days lost to sickness in the public sector is higher than in the private sector. The estimated coefficients for the middle and long distance variable are around 40% for the private sector sample than for the sample including public sector workers. This difference clarifies that the observed impact of commuting distance on absence is certainly not a public-sector phenomenon.

In a further robustness check, we exclude all observations for which sickness absence days amount to more than 30 days (model (iv)). Excluding outliers of the dependent variables may be important since workers who are absent more than 6 weeks may receive a wage reduction. The results for this restricted sample are virtually identical to those presented in Table 2, suggesting that outliers are not dominating the results. Hence, unobserved wage reductions due to long sickness absence do not affect our results.

Models (v), (vi) and (vii) are an attempt to test the sensitiveness of our results to reporting error by excluding observations that refer to small distance changes (less than 2 km) and by classifying commuting distance as a dummy variable (equals 1 if individual commutes more than 50 km) and a log-linear specification of commuting distance: In model (v) the effect of long distance commutes is very similar to the results reported in Table 2, whereas the coefficients of

short and middle distance commutes became insignificant. As expected, in model (vi) and (vii) the main effects are similar but smaller than that reported for the categorical commuting distance variable in Table 2.²⁵ In model (vi) the coefficient of the dummy variable indicate that the expected number of days absent is about 14% higher for those who travel more than 50 kilometres compared to those who travel less kilometres. In model (vii) the point estimate of the continuous commuting distance variable (in log), and therefore the elasticity, is 0.046 (s.e. 0.009). Thus, if the average logarithm of commuting distance, 2.24 in our data, falls about 0, sickness absence days will fall by about 10% (0.0467×2.24). van Ommeren and Gutiérrez-i-Puigarnau (2011) use a similar measure of commuting distance and find that commuting distance induces absence or shirking behaviour with an elasticity twice as large as the one we find (0.07 to 0.09). Since van Ommeren and Gutiérrez-i-Puigarnau (2011) also use data from the SOEP and a similar identification strategy, we can test whether it is the choice of the explanatory variables that is driving the difference in the results. To do this we re-run our analysis only using similar explanatory variables that were used by van Ommeren and Gutiérrez-i-Puigarnau (2011) (not reported). We find a point estimate of 0.062 indicating that failure to include additional confounders into the estimations is likely to result in overestimates of the strength of the association between commuting and sickness absence. We also re-ran the latter model using the categorical commuting distance variable instead of the continuous measure. We find that middle and long distance commutes are associated with higher sickness absence days, while short distance commutes are not. This illustrates that different measures of commuting distance may lead to different conclusions regarding the relevant variable. What is clear, however, is that the effect documented by van Ommeren and Gutiérrez-i-Puigarnau (2011) does not hold in general,

²⁵ Another attempt which we tried to deal with measurement error is to calculate a proxy for commuting time. To obtain the commuting time we built the difference between the daily working hours including travel time to and from work (taken from the question: “How many hours per normal workday do you spend on job, apprenticeship, second job (including travel time to and from work)?”) and the usual daily working hours (taken from the question: “And how many hours do you generally work, including any overtime?”) divided by 5 workdays. Again, we only find a positive and statistically significant effect of long commutes, particularly of commutes which take more than 45 minutes ($\beta_{46 \text{ min and more}} = 0.0559, p = 0.013$). We refuse to use this measure of commuting time as our focal explanatory variable as it is calculated in a very imprecise and inaccurate way. Further commuting time may be influenced by many reasons for instance by changes in congestion or infrastructure or even commuting modes.

as we find no evidence of an impact of short distance commutes on absence in our application. In this case the different findings may be due to the log-specification which does not allow for discontinuities in the impact of commuting on absence behaviour, as we observe.

The next four models ((viii) to (xi)) are an attempt to address a potential selection bias since considering a sample of workers who do not change employer and who do not change residence may create a selection bias. In our setting, there might be the following sources of endogeneity: Endogeneity might result from the self-selection of employees in a group of workers who do not change residence or employer. Strictly speaking we cannot exclude the possibility that individuals with unobserved positive attitudes to work are more likely to accept jobs at longer distances and are also less likely to be absent. Additionally, employees may, in fact, move residence or job as a reaction to employer-induced workplace relocation, which would be an endogenous reaction. Further, employees have some element of choice regarding a workplace move. If an employer needs some employees to move to a different part of the firm at a different location, employees are usually asked whether they are willing to move or not. When an employee is willing to do so, we deal with an endogenous move. To tackle the potential bias of this selection we have estimated the fixed-effects negative binomial model on other, less-selective, samples of data. So, we have included data on (viii) employees who change employer, (ix) employees who change residence, and (x) employees who change both employer and residence. Again we see that the effects of commuting distance reported in Table A.4, are similar to the results reported in Table 2, indicating that the effect of commuting long distances to work on sickness absence is not due to sample selection.

Furthermore, in the related literature that addresses commuting in general, many methods have been proposed to deal with problems of self-selection and endogeneity. Instrumental variables regression is considered to be an appropriate estimator in the presence of endogeneity, independent of what its possible cause may be. However, in the literature it is also emphasized that it is difficult to find valid instruments for commuting distance to correct for possible

endogeneity, as it is related to labour and residence locations behaviour (see e.g. Gutiérrez-i-Puigarnau and van Ommeren, 2010; Manning, 2003; Gubits, 2004). In order to properly address this issue we additionally apply the Hausman and Taylor (1981) instrumental variable estimation technique, which uses commuting distance from a period other than the current period as the instrument (model (xi)).²⁶ Accounting for endogeneity in this way does not seem to change our main results.

Overall, the robustness checks confirm the general finding that whereas sickness absence days are not affected by short commuting distances, sickness absence days are clearly higher for people who commute longer.

5.5 *Transmission Channels*

The previous analysis has uncovered a robust impact of commuting longer distances on the number of days of absence from work. In this subsection, we investigate various hypotheses concerning the cause of this relationship.

As outlined in the introduction, there is substantial evidence that commuting is associated with increased levels of illness. Since absence is negatively related to health (Puhani and Sonderhoff, 2010; Goerke and Pannenberg, 2015; for example, present according evidence for Germany), the impact of commuting on absence may be due to health effects. To rule out such impact, the estimations presented thus far, include a subjective measure of health. We further analyse health as transmission mechanism by, first, omitting the health variables included in the estimations depicted in Table 2. Second, we include additional health indicators, such as satisfaction with health, concern about individual's own health, degree of disability, the number of overnight hospital stays and the number of annual doctor visits.²⁷ Table 3 depicts the results.

²⁶ For detailed explanation, see Baltagi et al. (2007).

²⁷ We have also analysed the effect of interactions of distance with health indicators, but we did not find any significant effect, indicating that employee's marginal costs of commuting do not depend on individuals' state of health.

Table 3

Transmission channels. Fixed-effects estimates. Dependent variable: Days absent. Source: SOEP 2002 – 2011.

	Baseline (see Table 2)		Baseline without subjective health measures		Baseline with additional health measures	
	FE NEGBIN	FE OLS	FE NEGBIN	FE OLS	FE NEGBIN	FE OLS
Focal variable:						
Short distance	0.0442	1.305	0.0576*	1.363	0.0375	0.946
commuter	(1.91)	(1.94)	(2.49)	(1.89)	(1.62)	(1.42)
Middle distance	0.109***	2.607**	0.121***	2.809**	0.104***	2.339*
commuter	(3.65)	(2.79)	(4.04)	(2.91)	(3.46)	(2.55)
Long distance	0.191***	3.370*	0.205***	4.094**	0.195***	2.971*
commuter	(4.35)	(2.56)	(4.67)	(2.89)	(4.42)	(2.29)
Health status: very good (ref.)						
good	0.157***	1.242**			0.104**	0.255
	(4.67)	(3.18)			(3.02)	(0.66)
acceptable	0.343***	2.577***			0.200***	-0.323
	(9.52)	(4.93)			(5.10)	(-0.59)
less good	0.581***	10.87***			0.308***	4.422***
	(14.20)	(10.48)			(6.37)	(4.09)
bad	0.947***	47.50***			0.503***	35.68***
	(14.08)	(9.90)			(6.53)	(7.76)
Health satisfaction					-0.0414***	-0.947***
					(-6.67)	(-5.44)
Life satisfaction					0.00360	-0.285
					(0.56)	(-1.80)
Concerned about health: very (ref.)						
somewhat					-0.0179	-1.735*
					(-0.69)	(-2.20)
not at all					-0.0393	-1.253
					(-1.25)	(-1.50)
Invalidity level					0.000357	-0.172**
					(0.50)	(-2.77)
# of hospital stays					-0.00633***	-0.0828
					(-4.13)	(-1.16)
# of doctor visits					0.00802***	0.272***
					(18.21)	(9.63)
<i>N</i>	31,567	31,567	31,567	31,567	31,354	31,354

Notes: Only the coefficients for the commuting variables and those of potential health channels are reported. Non-commuters are treated as the reference category. The baseline models correspond to Model III and Model IV of Table 2. t statistics in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001.

Inspection of Table 3 reveals that more healthy people are indeed less absent from work. Moreover, we see that the magnitude of the estimated coefficients of the commuter variables decline if health indicators are included. This shows that part of the impact of commuting on absence is due to health reasons. However, the decline in the estimated coefficients is relatively moderate and they remain significant in the case of middle and long distance commuters. Therefore, we can conclude that commuting may deteriorate an individual's health but that this effect does not explain the observed impact of commuting on absence. This means that health is

not the channel through which commuting affects absence. If it were, then, when controlling for health, the significant coefficients on the commuting covariates would become statistically insignificant.

While health-related absence may be regarded as involuntary, the standard labour-supply perspective on absence views such behaviour as voluntary adjustment to predetermined and overly long working hours (Allen, 1981). Since commuting increases the length of the total workday while simultaneously reducing time for private use, the need to adjust total working time to the preferred amount is likely to be greater for individuals who commute. In order to scrutinise this transmission channel, we estimate extended specifications of equations (1) and (2), as depicted in Table 2, and add two dummy variables which indicate whether individuals would like to work less (or more) hours than they actually do. The estimated coefficients of the commuter variables are basically unaffected (relative to those depicted in Table 2). This is also true if we include further working time indicators, such as the number of actual hours worked, overtime hours per week or having a second job. Therefore, commuting does not result in greater voluntary absence.²⁸

In a substantial number of empirical studies, job (in-) security has been found to affect absence from work (see e.g. Staufenbiel et al., 2010; Bratberg et al., 2015). Moreover, reduced job security has a disciplining effect, suggesting that workers are more likely to accept jobs at longer distances and are also less likely to be absent. Hence, job insecurity may influence the probability of becoming a commuter and of being absent from work. We investigate this transmission channel by including a variable in extended specifications of equations (1) and (2) that indicates whether the individual is concerned about its own job security. Individuals who are not concerned about their job indeed have higher absence. However, the estimated coefficients of the commuting variables are basically the same as shown in Table 2. Alternatively, we have

²⁸ In a further step, we have also included information on private time use, for instance, the average time a day spent on running errands, housework, child care, care for people with disabilities and other dependants living in the household, leisure time, time for repairs and garden work. The estimated coefficients of the commuter variables are basically unaffected by the inclusion of the private time use variables.

employed the unemployment rate (at the level of federal states) as a proxy of job insecurity. Its inclusion does not substantially alter the estimated coefficients of interest. Consequently, the impact of commuting on absence does not arise because commuters are less concerned about their jobs.

We finally include a variable indicating whether individuals live in urban regions, rural regions or in regions undergoing urbanization. To include this information seems reasonable, since it cannot be excluded that employees living in rural regions are healthier and have a higher quality of life compared to individuals living in urban regions (Zeng et al., 2015). One may argue that commuting long distances is associated with living in rural regions and, hence, with a better quality of life which may weaken the effect of commuting on absence. Indeed, individuals who are living in rural regions have less sickness absence days. However, our results show that even after this additional variable is considered, the estimated coefficients of the commuter variables are basically unaffected (relative to those depicted in Table 2). The findings of the additional analyses are available from the authors upon request.

We conclude that long distance commuting raises absence, albeit not because commuting deteriorates health or alters job security, or because commuters face a greater mismatch between actual and desired working time. Alternatively, one may hypothesise that commuting is associated with lower work effort and, hence, more absenteeism. Our data does not enable us to provide further evidence on this kind of transmission mechanism. However, because the impact of commuting on absence can only be observed consistently for individuals who commute at least 25 km from home to workplace and is most pronounced for long distance commuters, this interpretation is questionable as well.²⁹

²⁹ One hypothesis we considered is that income, working hours or the desired working hours (work more or less hours) might be proxy indicators of work effort. Since the coefficient of long distance commutes is basically unaffected by the inclusion of these variables one may argue that the impact of commuting on absence does not arise because commuters provide lower work effort. However, as with other proxy indicators, there is a difficulty in ensuring that the claimed relationship is not confounded by other variables.

6. Conclusion

In summary, in this paper we enrich the literature on the relationship between commuting distance and sickness absence using panel data for Germany. Empirically, we know very little about this. We address possible reverse causality bias by using changes in commuting distance that are employer-driven and therefore exogenous. Specifically, we only exploit variation of commuting distance within individuals, when there are no changes in residence and employer. For Germany, we find a causal effect of commuting distance on sickness absence. This effect is apparent even when potential compensation in the housing and labour markets is taken into account. We find that absence from work is more prevalent among long distance commuters. Particularly long distance commutes imply an increase in the numbers of days absent of about 20% on average. The effect of middle distance commutes is much lower, i.e. about 11%. The effect becomes zero at commuting distances less than 25 kilometres. The results appear robust with specification and accounting for selection effects. We have explored some of the potential explanations for the effect of commutes to work on absence from work and suggest that it is not a result of working hours mismatch or a poor health. A deeper investigation of the determinants shows that differences in personnel characteristics, job related aspects and factors compensating for commuting are not able to explain the gap in sickness absence from work, and differences in health status only play a minor role. Taken together, these findings may point to potential shirking and moral hazard problems of long distance commuters.

Our findings have a number of implications. First, we demonstrate that sickness absence with respect to commuting is an important characteristic of the (German) labour market, which is in line with range of theoretical models (Zenou, 2002). Second, the present study suggests that commuting may have far-reaching consequences for both employees and the financial performance of the employing organisation. Hence, evidence of an absence-commute relationship puts a price on commuting distance and should be considered in cost-benefits assessments, since absence from work causes sizeable costs not only for the employer but also

for the employee. So, our findings point to the economic benefits from transport infrastructure improvements as well as to potentially saving costs to the health care system from sickness absence, at least if absent employees are entitled to sick pay. Third, the positive effects of commuting on sickness absence must be included in discussions on the expansion of economic regions and increasing the mobility of the workforce, a discussion that has been dominated by positive attitudes so far.

References

- Afsa, C., Givord, P., (2014). The impact of working conditions on sickness absence: a theoretical model and an empirical application to work schedules. *Empirical Economics*, Springer 46 (1), 285 – 305.
- Allen, S.G., (1981). An Empirical Model of Work Attendance. *The Review of Economics and Statistics* 63 (1), 77 – 87.
- Allison, P.D., (2009). Fixed effects regression models. *Series: Quantitative applications in the social sciences* 07 (160).
- Allison, P.D., Waterman, R.P., (2002). Fixed-effects negative binomial regression models. *Social Methodology* 32 (1), 247 – 265.
- Baltagi, B.H.; Bresson, G.; Pirotte, A. (2007). Fixed effects, random effects or Hausman-Taylor? A pretest estimator. *Economics Letters* 79 (3), 361 – 369.
- Becker, G.S., (1965). A theory of the allocation of time. *Economic Journal* 75 (299), 493 – 517.
- Besser, L.M., Marcus, M., Frumkin, H., (2008). Commute time and social capital in the US. *American Journal of Preventive Medicine* 34 (3), 207 – 211.
- Block, J., Goerke, L., Millán, J.M., Román, C. (2014). Family employees and absenteeism. *Economics Letters* 123 (2014), 94 – 99.
- Bratberg, E., Monstadt, K. (2015). Worried sick? Worker responses to a financial shock. *Labour Economics* 33 (2015), 111 – 120.
- Brown, S., Sessions, J.G., (1996). The Economics of Absence: Theory and Evidence. *Journal of Economic Surveys* 10 (1), 23 – 53.
- Cameron, C., Trivedi, P., (2008). *Regression analysis of count data*. Cambridge University Press.
- Carta, F., De Philippis, M., (2015). You've Come a Long Way, Baby. Effects of Commuting Times on Couples' Labour Supply. *Bank of Italy Temi di Discussione (Working Paper) No. 1003*.
- Costa, G., Pickup, L., Di Martino, V., (1988). Commuting – a further stress factor for working people: Evidence from the European Community. An empirical study. *International Archives of Occupational and Environmental Health* (1988) 60, 377 – 385.
- Cox, T., Griffiths, A., Rial-González, E., (2000). *Research on work-related stress*. Office for Official Publications of European Communities. Luxembourg.
- De Paola, M., (2010). Absenteeism and Peer Interaction Effects: Evidence from an Italian Public Institute. *The Journal of Socio-Economics* 39 (3), 420 – 428.
- Dionne, G., Dostie, B., (2007). New evidence on the determinants of absenteeism using linked employer-employee data. *Industrial and Labor Relations Review* 61 (1), 108 – 120.
- Evans, G.W., Wener, R.E., (2006). Rail commuting duration and passenger stress. *Health Psychology* 25(3), 408 – 412.
- Evans, G.W., Wener, R.E., (2007). Crowding and personal space invasion on the train: Please don't make me sit in the middle. *Journal of Environmental Psychology* 27 (1), 90 – 94.
- Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR), (2012). *Spatial Development Report 2011*. Bonn. Germany.

- Federal Statistical Office, (2008). Relocation of economic activities. Wiesbaden. Germany. (https://www.destatis.de/DE/Publikationen/Thematisch/UnternehmenHandwerk/VerlagerungAktivitaeten5529301069004.pdf?__blob=publicationFile).
- Federal Statistical Office, (2013): Census. Special microcensus evaluations: Commuting 2012. Wiesbaden. Germany.
- Flint, E., Cummins, S., Sacker, A., (2014). Associations between active commuting, body fat, and body mass index: population based, cross sectional study in the United Kingdom. *BMJ* 2014, 349:g4877.
- Frick, B., Malo, M.A., (2008). Labor market institutions and individual absenteeism in the European Union: the relative importance of sickness benefit systems and employment protection legislation. *Industrial Relations* 47 (4), 505 – 529.
- Gershenson, S., (2013). The causal effect of commute time on labor supply: Evidence from a natural experiment involving substitute teachers. *Transportation Research Part A: Policy and Practice* 54 (2013), 127 – 140.
- Gimenez-Nadal, J.I., Molina, J.A., (2014). Commuting time and labour supply in the Netherlands: A time use study. *Journal of Transport Economics and Policy* 48 (3), 409 – 426 (18).
- Goerke, L., Pannenberg, M., (2015). Trade union membership and sickness absence: Evidence from a sick pay reform. *Labour Economics* 33 (2015), 13 – 25.
- Gottholmseder, G., Nowotny, K., Pruckner, G.J., Theurl, E., (2009). Stress perception and commuting. *Health Economics* 18 (2009), 559 – 576.
- Gubits, D.B., (2004). Commuting, work hours, and the metropolitan labor supply gradient. Mimeo.
- Gulian, E., Matthews, G., Glendon, A.I., Davies, D.R., Debney, L.M., (1989). Dimensions of driver stress. *Ergonomics* 32 (6), 585 – 602.
- Gutiérrez-i-Puigarnau, E., van Ommeren, J.N., (2010). Labour supply and commuting. *Journal of Urban Economics* 68 (1), 82 – 89.
- Gutiérrez-i-Puigarnau, E., van Ommeren, J.N., (2013). Do rich households live farther away from their workplaces?. CPB Discussion Paper 244.
- Hansson, E., Mattisson, K., Björk, J., Östergren, P., Jakobsson, K., (2011). Relationship between commuting and health outcomes in a cross-sectional population survey in southern Sweden. *BMC Public Health* 2011, 11:843.
- Hilbe, J.M., (2007). Negative binomial regression. Cambridge University Press.
- Hennessy, D.A., (2008). The impact of commuter stress on workplace aggression. *Journal of Applied Social Psychology* 38 (9), 2315 – 2355.
- Hennessy, D.A., Wiesenthal, D.L., (1997). The relationship between traffic congestion, driver stress and direct versus indirect coping behaviours. *Ergonomics* 40 (3), 348 – 361.
- Jacobson, B.H., Aldana, S.G., Goetzel, R.Z., Vardell, K.D., Adams, T.B., Pietras, R.J., (1996). The relationship between perceived stress and self-reported illness-related absenteeism. *American Journal of Health Promotion* 11 (1), 54 – 61.
- Knox, J.B., (1961). Absenteeism and turnover in an Argentine factory. *American Sociological Review* 26 (3), 424 – 428.

- Koslowsky, M., Kluger, A., Reich, M., (1995). *Commuting stress. Causes, effects, and methods of coping.* Plenum Press. New York.
- Künn-Nelen, A., (2015). Does commuting affect health?. *Health Economics*, forthcoming, DOI: 10.1002/hec.3199.
- Leigh, J.P., (1983). Sex differences in absenteeism. *Industrial Relations* 22 (3), 349 – 361.
- Lindström, M., (2004). Social capital, the miniaturisation of community and self-reported global and psychological health. *Social Science & Medicine* 59 (3), 595 – 607.
- Livanos, I., Zangelides, A., (2013). Unemployment, labor market flexibility and absenteeism: a pan-european study. *Industrial Relations* 52 (2), 492 – 515.
- Lyons, G., Chatterjee, K., (2008). A Human Perspective on the Daily Commute: Costs, Benefits and Trade-offs. *Transportation Reviews* 28 (2), 181 – 198.
- Magee, C., Stefanic, N., Caputi, P., Iverson, D., (2011). Occupational factors and sick leave in Australian employees. *Journal of Occupational & Environmental Medicine* 53 (6), 627 – 632.
- Manning, A., (2003). The real thin theory: monopsony in modern labour markets. *Labour Economics* 10 (2), 105 – 131.
- Matthews, G., Dorn, L., Glendon, A.I., (1991). Personality correlates of driver stress. *Personality and Individual Differences*, 12, 53 – 549.
- Mattisson, K., Hakansson, C., Jakobsson, K. (2014). *Relationships between commuting and social capital among men and women in southern Sweden.* Environment and Behaviour 1 – 20. Sage Publications.
- National Travel Survey: 2012, (2013). National Travel Survey Statistics. Department for Transport. <https://www.gov.uk/government/collections/national-travel-survey-statistics>
- Novaco, R.W., Stokols, D., Campbell, J., Stokols, J., (1979). Transportation, stress, and community psychology. *American Journal of Community Psychology* 7 (4), 361 – 380.
- Novaco, R.W., Stokols, D., Milanesi, L. (1990). Objective and subjective dimensions of travel impedance as determinants of commuting stress. *American Journal of Community Psychology* 18 (2), 231 – 257.
- OECD, (2007). *OECD Rural Policy Review: Germany.* Paris. France.
- Olsson, L.E., Gärling, T., Ettema, D., Friman, M., Fujii, S., (2013). Happiness and satisfaction with work commute. *Social Indicators Research* 111 (1), 255 – 263.
- Puhani, P.A., Sonderhoff, K., (2010). The Effects of a Sick Pay Reform on Absence and on Health-Related Outcomes. *Journal of Health Economics* 29 (2), 285 – 302.
- Putnam, R.D., (2000). *Bowling alone: The collapse and revival of American community.* Touchstone. New York.
- Roberts, J., Hodgson, R., Dolan, P., (2011). “It’s driving her mad”: Gender differences in the effects of commuting on psychological health. *Journal of Health Economics* 30, 1064 – 1076.
- Rouwendal, J., (2004). Search theory and commuting behaviour. *Growth and Change* 35 (3), 391 – 417.

- Ross, S.L., Zenou, Y., (2008). Are shirking and leisure substitutable? An empirical test of efficiency wages based on urban economics theory. *Regional Science and Urban Economics* 38 (5), 498 – 517.
- Sandow, E., Westerlund, O., Lindgren, U., (2014). Is your commute killing you? On the Mortality risk of long-distance commuting. *Environment and planning A* 46(6), 1496 – 1516.
- Schaeffer, M.H., Street, S.W., Singer, J.E., Baum, A., (1988). Effects of control on the stress reactions of commuters. *Journal of Applied Social Psychology* 18 (11), 944 – 957.
- Schulze, S., (2009). Einige Beobachtungen zum Pendlerverhalten in Deutschland. Hamburg Institute of International Economics (HWWI) Policy Paper, No. 1-19.
- Staufenbiel, T., König, C. (2010). A model for the effects of job insecurity on performance, turnover intention and absenteeism. *Journal of Occupational and Organizational Psychology* 83 (1), 101 – 117.
- Stutzer, A., Frey, B.S., (2008). Stress that doesn't pay: The commuting paradox. *The Scandinavian Journal of Economics* 110 (2), 339 – 366.
- Taylor, P.J., Pocock, S.J., (1972). Commuter travel and sickness absence of London office workers. *British Journal of Preventive & Social Medicine* 26 (3), 165 – 172.
- Treble, J., Barmby, T., (2011). *Worker Absenteeism and Sick Pay*. Cambridge University Press. Cambridge et al.
- Wagner, G.G., Frick, J.R., Schupp, J., (2007). The German Socio-Economic Panel study (SOEP) – evolution, scope and enhancements. *Journal of Applied Social Science (Schmollers Jahrbuch)* 127 (1), 139–169.
- Wales, T.J., (1978). Labour supply and commuting time: An empirical study. *Journal of Econometrics* 8 (2), 215 – 226.
- Walsleben, J.A., Norman, R.G., Novak, R.D., O'Malley, E.B., Rapoport, D.M., Strohl, K.D., (1999). Sleep habits of Long Island rail road commuters. *Sleep* 22 (6), 728 – 734.
- Wener, R.E., Evans, G.W., Phillips, D., Nadler, N., (2003). Running for the 7:45: The effects of public transit improvements on commuter stress. *Transportation* 30 (2), 203 – 220.
- Wooldridge, J.M., (2010). *Econometric analysis of cross section and panel data*. MIT Press. Cambridge et al.
- VandenHeuvel, A.; Wooden, M. (1995). Do explanations of absenteeism differ for men and women?. *Human Relations* 48 (11), 1309 – 1329.
- van den Berg, G.J., Gorter, C., (1997). Job search and commuting time. *Journal of Business and Economic Statistics* 15 (2), 269 – 281.
- van Ommeren, J.N., (1998). On-the-job search behaviour: The importance of commuting time. *Land Economics* 74 (4), 526 – 540.
- van Ommeren, J.N., Gutiérrez-i-Puigarnau, E., (2011). Are workers with a long commute less productive? An empirical analysis of absenteeism. *Regional Science and Urban Economics* 41 (2011), 1 – 8.
- Zax, J.S., (1991). The substitution between moves and quits. *The Economic Journal* 101 (409), 1510 – 1521.

- Zax, J.S., Kain, J.F., (1996). Moving to the suburbs: Do relocating companies leave their black employees behind?. *Journal of Labor Economics* 14 (3), 472 – 504.
- Zeng, D., You, W., Mills, B., Alwang, J., Royster, M. Anson-Dwamena, R., (2015). A closer look at the rural health disparities: Insights from four major diseases in the Commonwealth of Virginia. *Social Science and Medicine* 140 (2015), 62 – 68.
- Zenou, Y., (2002). How do firms redline workers? *Journal of Urban Economics* 52 (3), 391 – 408.
- Zenou, Y., (2009). *Urban Labor Economics*. Cambridge University Press. Cambridge et al.
- Ziebarth, N.R., Karlsson, M., (2010). A natural experiment on sick pay cuts, sickness absence, and labor costs. *Journal of Public Economics* 94 (11 – 12), 1108 – 1122.

Appendix A

Table A.1

Variable definitions. Source: SOEP 2002 – 2011.

<i>Variable</i>	<i>Definition</i>
<i>Dependent variables</i>	
Days absent	Number of sickness absence days.
<i>Focal variable</i>	
Commuter	Commuting distance measured one-way in kilometres. Categorical Variable: 0 = “non-commuter (< 10 km)”, 1 = “short distance commuter (10 – 24 km)”, 2 = “middle distance commuter (25 – 49 km)”, 3 = “long distance commuter (> 49 km)”.
<i>Personal characteristics</i>	
Female	Dummy equals 1 for female. This variable is omitted in our fixed-effects regressions.
Age	Age in years.
Age ²	Age squared.
Married	Dummy equals 1 if the individual is either married or living as a couple.
Children	Dummy equals 1 if children in the household.
College Degree Education	Dummy equals 1 if individual has a completed college education.
Health status	A five point scale measuring highest level of education attainment: 0 = “no or other school certificate”, 1 = “sec. gen. school certificate”, 2 = “intermediate school degree”, 3 = “leaving certificate from voc. high school”, 4 = “college entrance exam”.
	A five point indicator of self-reported health status: 1 = “very good”, 2 = “good”, 3 = “acceptable”, 4 = “less good”, 5 = “bad”.
<i>Job related aspects</i>	
Tenure	Number of years in present job.
Tenure ²	Job tenure squared.
Working hours	Hours of work per week.
Regular part-time	Dummy equals 1 if individual works part-time.
Temporary job	Dummy equals 1 if individual has a temporary job.
Blue-collar worker	Dummy equals 1 if individual is a blue-collar worker.
Firm size	Size of company: 0 = “< 5 employees”, 1 = “5 – 19 employees”, 2 = “20 – 99 employees”, 3 = “100 – 199 employees”, 4 = “200 – 1999 employees”, 5 = “2000 employees and over”.
Public sector	Dummy equals 1 if individual works in the public sector.
Log (monthly wage)	Current gross labor income, being corrected by purchasing power parity and harmonized by consumer price index. Variable expressed in natural logarithms.
<i>Variables compensating for commuting</i>	
Satisfaction with work	Satisfaction with main job measured on an eleven point scale from 0 = “completely dissatisfied” to 10 = “completely satisfied”.
Satisfaction with leisure	Satisfaction with leisure time measured on an eleven point scale from 0 = “completely dissatisfied” to 10 = “completely satisfied”.
Satisfaction with dwelling	Satisfaction with dwelling measured on an eleven point scale from 0 = “completely dissatisfied” to 10 = “completely satisfied”.
HCI	Household crowding index defined as the number of usual residents in a dwelling divided by the number of rooms in the dwelling.
Industry	9 dummies equalling 1 for individuals working in the named industry: agriculture, energy, mining, manufacturing, construction, trade, transport, bank or insurance, services.
Region	Dummy variables for the Federal States of Germany: Schleswig-Holstein and Hamburg, Lower Saxony and Bremen, North Rhine-Westphalia, Hesse, Rhineland-Palatinate and Saarland, Baden-Wuerttemberg, Bavaria, Berlin, Brandenburg and Mecklenburg-Western Pomerania, Saxony, Saxony-Anhalt, Thuringia.
Year	Dummy variables for each year covered by the sample.

Table A.2

Descriptive statistics for full sample and for commuter categories. Source: SOEP 2002 – 2011.

	Full sample				Non-Commuter		Short distance commuter		Middle distance commuter		Long distance commuter	
	Mean	SD	Min.	Max.	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Number of days absent	10.36	24.70	0	365	9.96	24.31	10.59	25.12	10.43	23.72	11.86	27.62
Incidence of absence	0.6529	0.47	0	1	0.6328	0.48	0.6589	0.47	0.6786	0.46	0.7005	0.45
<i>Focal variable</i>												
Non-commuter (ref.)	0.46	0.49	0	1								
Short distance commuter	0.33	0.47	0	1								
Middle distance commuter	0.16	0.36	0	1								
Long distance commuter	0.05	0.23	0	1								
<i>Personal characteristics</i>												
Female	0.47	0.49	0	1	0.54	0.49	0.45	0.49	0.37	0.48	0.30	0.46
Age	45.25	9.00	19	64	45.64	9.00	45.08	9.01	44.71	8.82	44.87	9.34
Age ²	2129.56	797.12	361	4096	2164.33	802.61	2113.54	792.96	2076.96	774.96	2101.12	826.44
Married	0.74	0.43	0	1	0.73	0.44	0.75	0.43	0.74	0.43	0.76	0.42
Children	0.40	0.48	0	1	0.38	0.48	0.41	0.49	0.39	0.48	0.40	0.49
College degree	0.25	0.43	0	1	0.23	0.42	0.22	0.41	0.30	0.45	0.34	0.47
Education:	0.09	0.25	0	1	0.07	0.27	0.09	0.25	0.07	0.21	0.05	0.15
No school certificate (ref.)												
Sec. gen. school certificate	0.25	0.43	0	1	0.27	0.44	0.26	0.44	0.22	0.41	0.18	0.38
Intermediate school degree	0.38	0.48	0	1	0.39	0.48	0.37	0.47	0.38	0.48	0.40	0.49
Voc. high school	0.06	0.24	0	1	0.05	0.22	0.06	0.24	0.07	0.26	0.07	0.27
College entrance exam	0.22	0.41	0	1	0.19	0.39	0.22	0.41	0.26	0.43	0.30	0.45
Health status:	0.09	0.25	0	1	0.07	0.26	0.09	0.25	0.08	0.24	0.11	0.27
Very good (ref.)												
Good	0.46	0.49	0	1	0.47	0.49	0.45	0.49	0.45	0.49	0.44	0.49
Acceptable	0.34	0.47	0	1	0.33	0.47	0.35	0.47	0.36	0.48	0.33	0.47
Less good	0.10	0.30	0	1	0.09	0.29	0.10	0.31	0.10	0.30	0.11	0.31
Bad	0.01	0.10	0	1	0.01	0.11	0.01	0.10	0.01	0.10	0.01	0.12
<i>Job related aspects</i>												
Working hours	35.18	7.74	1.5	72.5	33.94	8.65	35.55	7.25	36.75	6.10	38.14	4.73
Regular part-time	0.21	0.40	0	1	0.27	0.44	0.19	0.39	0.12	0.33	0.06	0.25
Temporary job	0.02	0.15	0	1	0.02	0.14	0.02	0.15	0.02	0.14	0.03	0.17
Blue-collar worker	0.29	0.45	0	1	0.30	0.46	0.31	0.46	0.26	0.44	0.20	0.40
Firm size:	0.06	0.19	0	1	0.05	0.22	0.06	0.16	0.04	0.13	0.05	0.14
< 5 employees (ref.)												
5 – 19 employees	0.12	0.32	0	1	0.14	0.35	0.11	0.32	0.08	0.27	0.06	0.25

20 – 99 employees	0.19	0.39	0	1	0.22	0.41	0.18	0.39	0.15	0.36	0.15	0.36
100 – 199 employees	0.10	0.31	0	1	0.11	0.32	0.10	0.30	0.09	0.29	0.08	0.27
200 – 1999 employees	0.26	0.43	0	1	0.24	0.43	0.27	0.44	0.26	0.44	0.27	0.44
2000 employees and over	0.27	0.44	0	1	0.20	0.40	0.28	0.45	0.38	0.48	0.39	0.48
Public service	0.34	0.47	0	1	0.37	0.48	0.33	0.47	0.32	0.46	0.31	0.46
Tenure	14.95	9.52	2	49.80	15.06	9.60	15.23	9.49	14.45	9.32	13.51	9.48
Tenure ²	314.22	366.74	4	2480.04	319.22	375.89	322.19	362.93	299.60	349.89	272.52	359.54
Log (monthly wage)	7.76	0.57	4.23	10.23	7.62	0.60	7.80	0.53	7.94	0.51	8.06	0.47
<i>Variables compensating for commuting</i>												
Satisfaction with work	6.95	1.86	0	10	7.04	1.86	6.90	1.85	6.88	1.84	6.79	1.99
Satisfaction with leisure	6.61	2.01	0	10	6.77	1.98	6.56	1.99	6.46	1.97	5.99	2.19
Satisfaction with dwelling	7.82	1.72	0	10	7.80	1.73	7.84	1.68	7.83	1.70	7.80	1.81
HCI	0.70	0.28	0.09	10	0.71	0.29	0.70	0.28	0.67	0.28	0.69	0.27
N	31,567				14,113		10,435		5,129		1,890	

Table A.3

Estimation results. Dependent variable: Days absent. Source: SOEP 2002 – 2011.

	Model I Pooled NEGBIN	Model II Pooled OLS	Model III FE NEGBIN	Model IV FE OLS
<i>Focal variable</i>				
Non-commuter (ref.)				
Short distance commuter	0.0385 (1.45)	0.572 (1.86)	0.0442 (1.91)	1.305 (1.94)
Middle distance commuter	0.0705* (2.17)	0.846* (2.17)	0.109*** (3.65)	2.607** (2.79)
Long distance commuter	0.201*** (3.72)	2.173*** (3.36)	0.191*** (4.35)	3.370* (2.56)
<i>Personal characteristics</i>				
Female	0.191*** (6.52)	2.073*** (6.11)	0.256*** (8.77)	
Age	-0.0362*** (-3.33)	-0.374** (-2.96)	-0.0976*** (-9.11)	-1.917*** (-5.09)
Age ²	0.000525*** (4.20)	0.00565*** (3.69)	0.000829*** (6.92)	0.0234*** (5.39)
Married	0.0133 (0.44)	0.0644 (0.17)	0.00987 (0.37)	0.654 (0.62)
Children	-0.0561 (-1.91)	0.0875 (0.27)	0.0654** (2.75)	0.326 (0.48)
College degree	-0.181*** (-5.01)	-1.245** (-3.18)	-0.0259 (-0.69)	-1.622 (-0.95)
Education: No school certificate (ref.)				
Sec. gen. school certificate	0.0660 (1.36)	0.454 (0.67)	-0.00179 (-0.04)	7.260*** (5.86)
Intermediate school degree	-0.0864 (-1.78)	-0.535 (-0.81)	0.129* (2.53)	3.153** (2.85)
Voc. high school	-0.0722 (-1.08)	-1.000 (-1.30)	0.239*** (3.56)	0.932 (0.60)
College entrance exam	-0.151** (-2.73)	-1.231 (-1.83)	0.322*** (5.55)	5.702*** (4.60)
Health status: Very good (ref.)				
Good	0.270*** (5.80)	1.386*** (5.33)	0.157*** (4.67)	1.242** (3.18)
Acceptable	0.661*** (13.47)	4.557*** (13.36)	0.343*** (9.52)	2.577*** (4.93)
Less good	1.335*** (23.28)	15.65*** (19.74)	0.581*** (14.20)	10.87*** (10.48)
Bad	2.381*** (26.99)	58.93*** (13.43)	0.947*** (14.08)	47.50*** (9.90)
<i>Job related aspects</i>				
Working hours	0.0173*** (6.29)	0.165*** (5.30)	0.00776*** (3.67)	0.0736 (1.18)
Regular part-time	0.149** (3.23)	1.244* (2.12)	0.0462 (1.30)	-0.555 (-0.61)
Temporary job	-0.101 (-1.39)	-1.471* (-2.43)	-0.0922 (-1.59)	-2.156* (-2.39)
Blue-collar worker	0.286*** (8.36)	2.787*** (6.67)	-0.0338 (-1.18)	1.350 (1.39)
Firm size: < 5 employees (ref.)				
5 – 19 employees	0.127 (1.64)	0.987 (1.55)	0.249*** (3.99)	-0.835 (-0.67)
20 – 99 employees	0.293*** (3.87)	2.608*** (4.00)	0.371*** (5.95)	2.459 (1.62)
100 – 199 employees	0.366*** (4.52)	3.304*** (4.60)	0.383*** (5.87)	2.338 (1.47)
200 – 1999 employees	0.444*** (5.87)	4.428*** (6.76)	0.454*** (7.25)	3.626* (2.45)
2000 employees and over	0.468***	4.578***	0.446***	2.700

	(6.15)	(6.90)	(7.10)	(1.80)
Public service	0.158***	1.487***	0.177***	0.388
	(4.76)	(3.98)	(6.31)	(0.50)
Tenure	-0.00542	0.0257	-0.000529	0.462**
	(-1.15)	(0.46)	(-0.13)	(3.21)
Tenure ²	0.0000772	-0.00109	0.0000895	-0.00661
	(0.64)	(-0.69)	(0.84)	(-1.61)
Log (monthly wage)	-0.0724	-1.107*	0.184***	-1.344
	(-1.88)	(-2.58)	(6.43)	(-1.44)
<i>Variables compensating for commuting</i>				
Satisfaction with work	-0.0496***	-0.566***	-0.0341***	-0.401**
	(-7.11)	(-5.46)	(-7.29)	(-2.69)
Satisfaction with leisure	0.0138*	0.263***	0.0139**	0.202
	(2.25)	(3.35)	(3.01)	(1.71)
Satisfaction with dwelling	0.00636	0.176	-0.00291	-0.0499
	(0.84)	(1.80)	(-0.53)	(-0.37)
HCI	-0.0331	-0.418	-0.0826*	-0.926
	(-0.70)	(-0.76)	(-2.08)	(-0.96)
Business sector dummies	Included	Included	Included	Included
Region dummies	Included	Included	Included	Included
Year dummies	Included	Included	Included	Included
_cons	1.962***	10.48*	-0.893**	51.10***
	(5.23)	(2.38)	(-2.72)	(4.22)
Inalpha				
_cons	1.005***			
	(89.23)			
N	31,567	31,567	31,567	31,567

Notes: *t* statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.4

Robustness checks. FE NEGBIN estimates. Dependent variable: Days absent. Source: SOEP 2002 – 2011.

		N	Focal variable	FE NEGBIN
Sample used for Table 2		31,567	Short distance commuter	0.0442 (1.91)
			Middle distance commuter	0.109*** (3.65)
			Long distance commuter	0.191*** (4.35)
(i) Women		14,942	Short distance commuter	0.0366 (1.12)
			Middle distance commuter	0.1208** (2.65)
			Long distance commuter	0.162* (2.19)
(ii) Men		16,625	Short distance commuter	0.0555 (1.68)
			Middle distance commuter	0.102* (2.54)
			Long distance commuter	0.212*** (3.84)
(iii) Sample excluding public sector		20,279	Short distance commuter	0.0277 (0.93)
			Middle distance commuter	0.1537*** (4.04)
			Long distance commuter	0.2586*** (4.66)
(iv) Sample excluding 'sickness absence days outliers'		29,838	Short distance commuter	0.0418 (1.73)
			Middle distance commuter	0.0967*** (3.10)
			Long distance commuter	0.1667*** (3.66)
(v) Sample excluding small distance changes		22,117	Short distance commuter	0.0181 (0.63)
			Middle distance commuter	0.0454 (1.28)
			Long distance commuter	0.1609*** (3.31)
(vi)	Commuting distance as dummy variable	31,567	Commutes: 0 – 49 km (ref). Commutes: 50 km and more	0.1423*** (3.44)
(vii)	Commuting distance as log-linear specification	31,190	Commuting distance (in log)	0.0467*** (4.96)

(viii)	Sample including employer change	32,178	Short distance commuter	0.0398 (1.74)
			Middle distance commuter	0.110 ^{***} (3.72)
			Long distance commuter	0.181 ^{***} (4.24)
(ix)	Sample including residence change	39,181	Short distance commuter	0.0381 (1.90)
			Middle distance commuter	0.0884 ^{***} (3.41)
			Long distance commuter	0.177 ^{***} (4.74)
(x)	Sample including employer and residence change	40,289	Short distance commuter	0.0360 (1.84)
			Middle distance commuter	0.0927 ^{***} (3.66)
			Long distance commuter	0.175 ^{***} (4.84)
(xi)	Hausman-Taylor IV estimation	31,567	Short distance commuter	1.460 [*] (2.08)
			Middle distance commuter	2.969 ^{**} (3.08)
			Long distance commuter	4.418 ^{**} (3.22)

Notes: Only the coefficients for the commuting variables are reported. In models (i) – (v) and models (viii) – (xi) non-commuters are treated as the reference category. Like in the main table, all control variables are included in all specifications. t statistics in parentheses. *p<0.05, **p<0.01, ***p<0.001.

Figure A.1
Distribution of absence days. Source: SOEP 2002 – 2011.

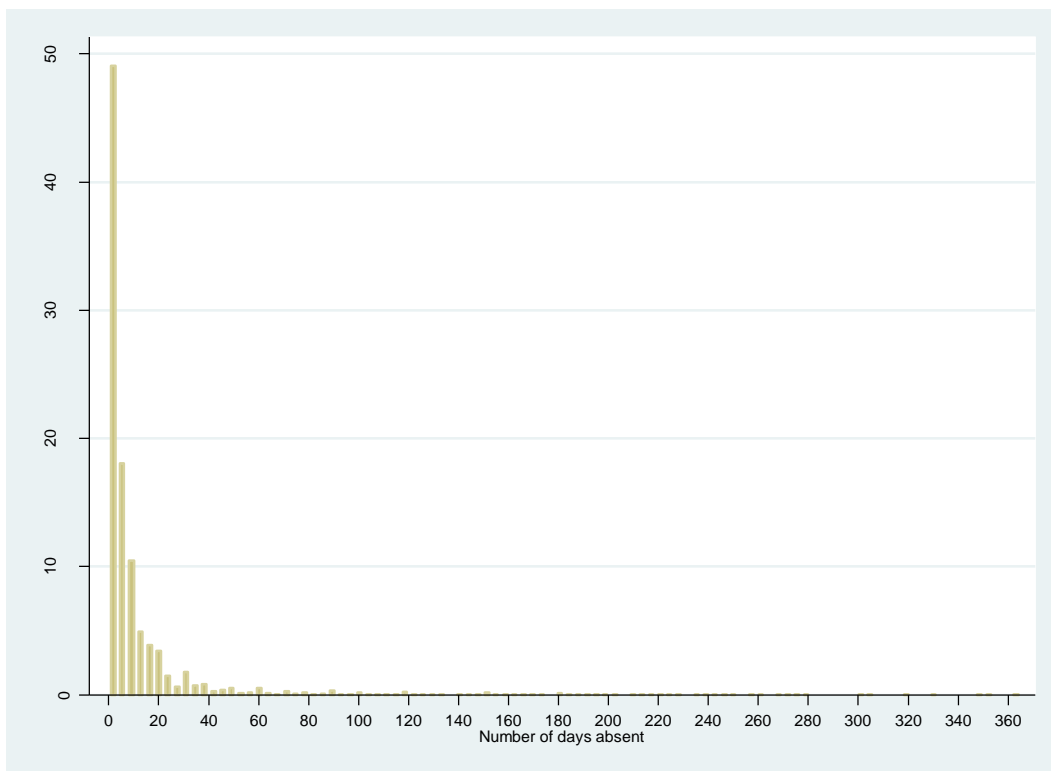


Figure A.2
Distribution of commuting distances. Source: SOEP 2002 – 2011.

