IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11700

Commuting Time and Sick-Day Absence of US Workers

Jose I. Gimenez-Nadal
Jose Alberto Molina
Jorge Velilla

JULY 2018

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11700

Commuting Time and Sick-Day Absence of US Workers

Jose I. Gimenez-Nadal
University of Zaragoza and BIFI

Jose Alberto Molina
University of Zaragoza, BIFI and IZA

Jorge Velilla
University of Zaragoza

JULY 2018

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Commuting Time and Sick-Day Absence of US Workers*

This paper analyzes the relationship between commuting time and sick-day absence of US workers. Using data from the Panel Study of Income Dynamics for the years 2011, 2013, and 2015, we find that a 1% increase in the daily commute of male workers is associated with an increase of around 0.018% in sick-day absences per year. In the case of women, the relationship is not significant. These results hold after controlling for individual fixed effects and socio-demographic characteristics, changes in jobs and places of residence, and differences in the self-reported health status of workers. By determining how commuting time is related to sickness absenteeism, we shed light on the relationship between commuting behavior and workers' health-related outcomes, measured by their labour supply.

JEL Classification: I10, J22, R2, R40

Keywords: commuting time, sickness absence, health-related outcomes, labour supply

Corresponding author:
Jose I. Gimenez-Nadal
Department of Economic Analysis
Faculty of Economic and Business Studies
University of Zaragoza
C. Gran Via 2
50005 Zaragoza
Spain
E-mail: ngimenez@unizar.es

* This paper has benefited from funding from the "Programa Operativo FSE Aragón 2014-2020" and from the Regional Government of Aragon and the European Fund of Regional Development (Grant S32_17R).

1. Introduction

In this paper, we analyze the relationship between commuting time and sick-day absences of workers in the US. The analysis of commuting has gained much relevance in the literature in recent decades (see Ma and Banister (2006) for a chronological review), as a result of increases in the time/distance workers in developed countries devote to commuting to/from work (Kirby and LeSage, 2009; Gimenez-Nadal and Molina, 2014; 2016; Goerke and Lorenz, 2017). Commuting has been linked to negative health-related outcomes (Hansson et al., 2011; Künn-Nelen, 2016), which include lower subjective/psychological well-being (Roberts et al., 2011; Dickerson, Hole and Munford, 2014) and increased stress (Gottholmseder et al., 2009; Wener et al., 2003; Frey and Stutzer, 2008; Novaco and Gonzalez, 2009). The negative effects of commuting have also been linked to labor costs (Allen, 1983) and losses in productivity (Grinza and Rycx, 2018), in that commuting is related to increased absence of workers due to sickness (van Ommeren and Gutierrez-i-Puigarnau, 2011).

Within this framework, the goal of this paper is to provide empirical evidence of the relationship between workers' commuting and sick-day absence. The literature on this topic is quite scarce as, to the best of our knowledge, only van Ommeren and Gutierrez-i-Puigarnau (2011) and Goerke and Lorenz (2017) have previously studied the relationship. Their results refer to Germany, and their conclusions are contradictory. To bridge the gap in the literature, we use data from the Panel Study of Income Dynamics of the United States, for the years 2011, 2013, and 2015, and analyze the relationship between workers' commuting time and sick-day absence.

We find that male workers who commute longer are more likely to be absent from work due to sickness. In particular, we find that the elasticity between daily commuting time and the annual sick-day absences is estimated to be around 0.018. These results are robust to considering individual fixed effects and socio-demographic characteristics, changes in jobs and places of residence, and differences in workers' self-reported health status. We find no significant relationship for US female workers. We consider that this gender difference may be due to females' shorter commuting times, given that they have greater responsibilities for day-to-day household tasks in comparison to males, including housework and childcare (Roberts et al., 2011; Gimenez-Nadal and Molina, 2016).

Our contribution to the literature is twofold. First, we complement prior analyses from van Ommeren and Gutierrez-i-Puigarnau (2011) and Goerke and Lorenz (2017), adding evidence of the relationship between (long) commuting and sick-day absence of workers. This paper presents the first estimate of this relationship for the case of the United States, and sheds light on the relationship between commuting behavior and health-related outcomes, suggesting that employees with longer commutes are more likely to be absent from work due to sickness. Second, sick-day absence may be considered a (negative) measure of worker performance and productivity. Thus, we contribute to the literature focusing on the relationship between commuting time and worker productivity, usually measured by hours of labour supply (Gutierrez-and-Puigarnau and van Ommeren, 2010; 2015; Gershenson, 2013; Gimenez-Nadal and Molina, 2014; 2016; Carta and De Philippis, 2018). The results presented in this paper may be complementary to the latter field of research.

The rest of the paper is organized as follows. Section 2 presents the data and the variables used in the empirical analysis. Section 3 shows the econometric strategy, and Section 4 shows the main results. Finally, Section 5 sets out the main conclusions.

2. Data and variables

We use data from the Panel Study of Income Dynamics (PSID). The PSID is a longitudinal household survey, conducted every two years by the University of Michigan, since 1986 (https://psidonline.isr.umich.edu/). The PSID consists of a representative sample of more than 5,000 US households per wave, and contains information on several factors, including socio-demographics, employment, and wealth. The PSID includes information at the household level, and on every member of the interviewed households. Waves of the PSID before 2011 cannot be used throughout the analysis, as information about commuting time was first included only in2011. Thus, we use data from the 2011, 2013, and 2015 waves of the PSID interviews.

From each household, we select individuals who are defined as the head of the household and the spouse (if any), and restrict the sample to employed individuals who report positive labor supply, and thus students, retired workers, and disabled workers are omitted from the analysis.[1] Self-employed workers are also excluded from the analysis, as their commuting routines may be of a different nature than those of employees (van Ommeren and van der Straaten, 2008; Gimenez-Nadal et al., 2018a), so the relationship between commuting and sick-day absences will be different. Individuals who appear in the sample for just one year are also omitted, as we aim to take advantage of the panel structure of the data to net out this relationship from the unobserved heterogeneity of individuals. We also identify and omit outliers from the analysis, to avoid sensitivity problems in the econometric analysis.[2] These restrictions leave us with an unbalanced panel of 14,894 observations, that correspond to 6,010 individuals (the average individual appears in the sample 2.478 times), of whom 3,306 are men and 2,704 are women.

The PSID contains information about certain sources of absence during the previous year, including (own) sickness, absence because another person in the household was sick, strikes, or vacations and time off, measured as the number of work days individuals missed. Strikes, vacations, and time off cannot be defined as a form of sick-day absence, and thus they are not considered in the analysis. Since the potential effects of a worker's commute on his/her spouse's sick-day absences is not within the scope of our analysis, we define sick-day absence only as days of absence from work due to own sickness. Table 1 shows summary statistics of such absences, for men and women. We observe that men are absent from work, on average, 0.78 days per year because of sickness, vs 0.74 days in the case of women. The gender difference in sick-day absence is not statistically significant at the 95 percent level of confidence, according to a *t*-type test.

Our main explanatory variable is the time devoted to commuting to/from work. This information was first collected in the PSID in the year 2011, and is measured in minutes per

[1] Including employed workers with positive labor supply may lead to sample selection bias. However, as shown in Section 4, estimates are robust to sample selection issues.

[2] We detect outliers using the Blocked Adaptive Computationally Efficient Outlier Nominators (BACON) algorithm (Billor et al., 2000). We find only one outlier, which corresponds to an individual who reported 208 sick-day absences. This observation is not considered in the analysis.

day.[3] It is important to note that commuting times are usually studied using time-use surveys (e.g., Gimenez-Nadal and Molina, 2016; Gimenez-Nadal et al., 2018b), as time-use data based on diaries is more accurate than time-use data based on stylized questionnaires (Juster and Stafford, 1985; Bianchi et al., 2000; Bonke, 2005; Yee-Kan, 2008). Conversely, information collected from diaries, such as the ATUS, depends on the diary day, which is sometimes a day with unusual schedules (e.g., "day" bias). Against information based on time-use surveys, data collected in the PSID refers to minutes of commuting to and from work on a "typical day", and hence does not suffer from this "day" bias. Table 1 shows that male workers spend, on average, 43.20 minutes per day in commuting, against 37.53 average minutes spent by women.[4] This difference of 15.11% is statistically significant at standard levels, in line with prior studies addressing the gender gap in commuting (e.g., Gimenez-Nadal and Molina, 2014, 2016).

The PSID provides information about individual, family, and labor characteristics that may be correlated with sick-day absences, and that will be considered as explanatory variables in the econometric models. We consider age as a factor that may affect the number of sick-day absences, given that as people get older their health status may worsen, leading to an increase in the number of such absences. Education is also considered, as prior research has highlighted the negative educational gradient in sick-day absence (Hämmig and Bauer, 2013; Kaikkonen et al., 2015; Piha et al., 2009; Piha et al., 2013). Education is collected in the PSID as the "highest grade or year of schooling completed", and measured in completed years of education.

We also control for the number of hours worked per week (i.e., in a "typical week"), the annual salary, measured in dollars, and total family income, defined as all the income

[3] Information is obtained from the PSID with the following question: "On a typical day, how many minutes is (was) your round trip commute to and from work?". Time is in general more accurate than distance to measure commutes, which leads to a reduced error term (Small and Song, 1992; van Ommeren and van der Straaten, 2008; Jara-Díaz and Rosales-Salas, 2015). Furthermore, times directly collect a series of aspects of importance, such as traffic density, or speed of commutes, that distances do not usually capture (Gimenez-Nadal et al., 2018a).

[4] Only 5.4% of the sample reports zero commuting (6.1% of males, and 4.5% of females). The standard deviation of being a zero commuter within individuals is small (0.12 against a standard deviation of 0.23 when considering cross-sectional sample). This indicates that not commuting to/from work is relatively persistent within workers. Zero commuters are not excluded from the analysis, as omitting zero commuters reduces the variability and may lead to sample selection issues. In any case, results are robust to their exclusion as will be shown in Section 4.

received by the household, and measured in dollars per year, which allows us to control for the socio-economic position of the individual.[5] Better socio-economic positions have been linked to lower sick-day absence (Barmby et al., 1994, 1995; Piha et al., 2009; MarKussen et al., 2011; Löve et al., 2013). Furthermore, there may be differences in sick-day absence depending on the ethnic status of the worker (Baker and Pocock, 1982; Leigh, 1991). Thus, we create a dummy variable that takes value "1" if the individual is white, and value "0" otherwise. We also consider the self-reported health status of individuals at the time of the interview. The PSID allows us to define five levels of (self-reported) health, and five dummies are defined: excellent, very good, good, fair, and poor. It is important to control for health, to isolate the effect of potential shocks in health on sick-ay absenteeism, from the effects of commuting (Leigh, 1991; van Ommeren and Gutierrez-i-Puigarnau, 2011).

Household characteristics, such as the presence/age of children, and marital status, have been linked to sick-day absence (Mastekaasa, 2000; Bratberg and Naz, 2009; Simonsen and Skipper, 2012). For instance, Mastekaasa (2000) argues that women are more absent from work due to sickness because women, to a greater extent than men, are exposed to the 'double burden' of combining paid work with family obligations, particularly for married women. Thus, the presence of children and the marital status of the couple may condition the number of days of sick-day absence. For the number of children, we follow Campaña et al. (2016) and consider the number of children in two age-groups, children under 7 years, and children between 7 and 17 years (inclusive). For the two age groups, we create dummy variables that take value "1" if there are one or more children in the household at this age range, and value "0" otherwise. Regarding marital status, we create a dummy variable that takes value "1" if the corresponding individual lives with a (married or unmarried) partner, and "0" otherwise.

Finally, to control for changes in work and/or residence locations, we include a dummy variable that takes value 1 if individuals have moved to another residence in the year prior

[5] An individuals' annual salary is defined as the total gross labor income, measured in dollars per year, received the previous year ("How much did you earn altogether from wages or salaries in [the previous year], that is, before anything was deducted for taxes and other things?"). The total income received by households, and measured in dollars per year, refers to the sum of all the sources of income of households, including taxable income of all family members, transfer income of all family members, and Social Security income of all family members.

to the interview (and value “0” otherwise) and an analogous variable for whether individuals changed their jobs. By including these controls, we aim to isolate shocks in sick-day absence due to relocations of workers within firms or to other residential locations (van Ommeren and Gutierrez-i-Puigarnau, 2011), and also by changes in the urban structure of cities and metropolitan areas.[6] Table 1 shows that 6.66% of the analyzed sample changed to a new job (7.6% of the males, and 5.4% of the females) during the analyzed period, but in 30.17% of the observations (33.7% of the men, and 25.3% of the women) individuals report having changed their residence. The fact that we can control for these changes will allow us to obtain the relationship between commuting time and sick-day absence, net of changes in sick-day absence that are due to changes in job and residential location.

3. Econometric analysis

The empirical strategy is based on the identification strategy of van Ommeren and Guiterrez-i-Puigarnau (2011) using information on commuting distances, from the GSOEP. We exploit the panel structure of the data, and link changes in the commuting time of workers to changes in sick-day absence of those same workers, controlling for a set of cofounders, and using worker, job, and residence fixed effects. It is important to note that, in the reference work of van Ommeren and Guiterrez-i-Puigarnau (2011), individuals whose commute changes during the interview year are removed. However, this restriction does not affect the main results (see Appendix C in van Ommeren and Gutierrez-i-Puigarnau, 2011).

We estimate the following linear Fixed Effects (FE) model:[7]

$$\log(Y_{it}) = \alpha_i + \beta_C \log C_{it} + \beta_j X_{it} + u_{it}, \qquad (1)$$

where α_i represents the unobserved time-invariant effect of individual “i”, Y_{it} represents the sick-day absences of individual “i” in wave “t” (t=2011, 2013, 2015), C_{it} is the daily minutes of commuting of individual “i” in wave “t”, X_{it} represents a vector of the socio-

[6] See Gimenez-Nadal et al. (2018a) for a recent review of the importance of urban structures in commuting trips of US workers.

[7] A Hausman test p-value of 0.015 rejects the random effects estimator, against the fixed effects estimator.

demographic characteristics we control for individual "i" in wave "t", and u_{it} is the error term. We also include region dummies at the state level, year fixed effects, and occupation fixed effects.[8] Further, as men and women tend to show different behaviors in their time-allocation decisions, we estimate Equation (1) separately for male and female workers (Gimenez-Nadal and Sevilla, 2011; 2012). We transform commuting time and sick-day absences to their log form so that β_C can be interpreted in terms of elasticity: the percent change in y (the dependent variable), while x (the independent variable) increases by one percent. Figure 1 shows k-density functions of the log of commuting time, and we observe that the distribution of commuting time is very similar between men and women, with a peak at zero (i.e., non-commuters) and an inverted u-shaped distribution that resembles the shape of a normal distribution.

We first estimate a baseline model excluding controls for health status and changes in job and residence location, and we then estimate a second model where we include these factors as explanatory variables. Regarding the self-reported health status of individuals, the negative effect of long commutes on health is a widely-studied topic (e.g., Novaco et al., 1990; Gottholmseder et al., 2009; Hansson et al., 2011; Roberts et al., 2011), and it is likely that a meaningful part of the sick-day absence behaviors of workers is motivated by differences in their general health status. On the other hand, as the health index used is subjective, there may be unobserved factors affecting both health and sick-day absence that are biasing the results. Moreover, the relationship between commuting and sick-day absence may be affected by the same unobserved factors that affect changes in job and residence locations, and thus, in the baseline model, we exclude controls for changes in job and/or residence location and then include them in the full model. If estimates are insensitive to the inclusion of the health status and changes in job and residence locations, this would suggest that these unobserved factors have a non-significant effect on the

[8] Despite that theoretical models tend to consider workers as being homogeneous, the empirical evidence has shown that workers in different occupations show different commuting behaviors (Hanson and Johnston, 1985; Gordon et al., 1989; Hanson and Pratt, 1995). The PSID identifies 456 occupations, which are aggregated in 25 groups, according to the 2000 Census of Population and Housing: Alphabetical Index of Industries and Occupations. For simplicity, we define the following 14 types of occupation: Management; Science and technology; Services; Arts; Health; Catering; Maintenance; Sales; Farming, fishing and forestry; Construction; Installation; Production; Transport; and Other occupations.

estimated relationship.[9] However, we must note that van Ommeren and Gutierrez-i-Puigarnau (2011) argue that the full model may face some endogeneity issues, related to workers' choices about workplace movements within the same job, as voluntary firm relocations or promotions. These relocations may have an impact on workers' commutes. However, these relocations are mostly measured by controlling for changes in jobs, so the endogeneity issue is minimized.

4. Results

Table 2 shows the results of estimating Equation (1) for the (log of) sick-day absences. Column (1) shows the results when men and women are considered together, and we repeat estimates separately for men (Column (2)) and women (Column (3)). Columns (4), (5), and (6) are analogous but include controls for health indicators, and changes in job and residence location. First, we observe that estimates of parameters are qualitatively invariant to the inclusion of the latter controls, indicating that time-varyiant unobserved factors affecting the relationship between sick-day absence and commuting time are different to those related to self-reported health status, and changes in job and residence location. Regarding the relationship between commuting time and sick-day absence, we observe a positive and statistically-significant association between commuting time and the annual sick-day absences in general terms, with a one-percent increase in the daily minutes spent commuting leading to an increase of 0.016% in sick-day absences. But when the analysis is done separately by gender, we observe that the association is concentrated on male workers, with an elasticity of 0.017%, while in the case of women the effect is found to be non-statistically significant.

These results confirm the estimates of van Ommeren and Gutierrez-i-Puigarnau (2011), and are novel as they refer to the United States. In addition, our empirical analysis finds important gender differences, as in the case of women the relationship is found to be non-statistically significant. The empirical analysis provided by van Ommeren and Gutierrez-i-Puigarnau (2011) is not conducted by gender, and only as a sensitivity check is it found that

[9] Given that we use FE linear models, we refer to time-variant unobserved factors, given that time-invariant unobserved factors are controlled for with the FE estimator.

the effect is higher for men, but still significant for women. Gender differences can be explained in several ways, such as differences in the type of jobs men and women occupy, or social norms. Regarding the latter, women may be supposed to be in charge of the household responsibilities, which affects the commuting behavior of women (see the Household Responsibilities Hypothesis (HRH), whose effects on commuting have been found to be significant, in Gimenez-Nadal and Molina (2016)). Given the household responsibilities of women, women usually assume most of the household production, which includes unpaid work and childcare, leading to gender differences in the time devoted to paid and unpaid work in the US (Aguiar and Hurst, 2007). This could explain why female workers have shorter commutes than men, and thus commuting becomes non-significant in the explanation of sick-day absences in favour of other uses of time, such as childcare and unpaid work. Further research into possible explanations for this gender difference is needed.

Regarding the rest of the explanatory variables, in general, very few socio-demographic characteristics are related to sick-day absence. For instance, only the number of hours worked per week shows a positive and significant effect on sick-day absence (and only in the general case). In the case of female workers, income shows an inverted u-shaped, statistically-significant correlation with sick-day absence. The health indicators show a significant relationship in general terms. In comparison with those with excellent (self-reported) health, workers with a "fair" health status show a higher propensity to show absenteeism behavior, and this relationship is concentrated for male workers. Finally, changes in residence location show a non-statistically-significant effect, but changes in job shows a statistically-significant and negative relationship to sick-day absence in the case of male workers. This may be due to the need to make a good "first impression" when starting in a new job or a new position, which reduces the likelihood of a male worker being absent from the job due to sickness.

We now develop additional analyses to check the robustness of our results.[10] First, to minimize the effect of atypical workers, we eliminate workers with more than 15 sick-day absences (0.03% of the sample), and more than 120 minutes of commuting (2.10% of the

[10] We only show results for the log of commuting time. See Table A1 in the Appendix for the full set of results.

sample). Second, we propose a fixed-effects negative binomial specification, as the dependent variable takes integer values (which only in this case is not measured in logarithms).[11] Third, we exclude zero-commuters from the analysis. Fourth, we estimate the model including self-employed workers. It is important to note that self-employed workers do not receive wages, and the costs of sick-day absence are normally assumed by them, so factors affecting such absences may differ between employed and self-employed workers.

Table 3 shows estimates of the parameter of interest. In all of the different specifications, estimates are robust to Table 2. In general terms, we find a positive and statistically significant effect of longer commutes on sick-day absences, which can be exclusively attributed to men. In the case of women, the parameters are non-significant at standard levels. For instance, the model including only commuter workers indicates that, by excluding zero-commuters, the estimated effect would be higher than in general terms (although the estimated standard error increases, and then the parameters are significant only at the 10% level). Estimates of the model including the self-employed are also robust. This implies that, even when self-employed workers may have a different behavior from employees, the relationship between commuting time and sick-day absence is robust to the consideration of self-employed workers.

5. Conclusions

This paper provides empirical evidence of the relationship between workers' commuting time and sick-day absence, to shed light on the relationship between commuting behavior and health-related outcomes. To the extent that workers spend a non-negligible part of their working days going to and from work, and that it may affect workers' productivity, the analysis of this topic is important. Using data from the 2011, 2013, and 2015 waves of the Panel Study of Income Dynamics of the United States, we show a positive and significant

[11] The negative binomial model is typical in the literature, and should be more convenient to model count variables, as is the case of annual sick-day absences (for instance, conditional and unconditional fixed-effect negative binomial estimates are used in van Ommeren and Gutierrez-i-Puigarnau (2011) to find that conditional models show downwardly minimal biases). However, conditional fixed-effects negative binomial models have been criticized, and unconditional negative binomial models tend to underestimate error terms (Allison and Waterman, 2002; Greene, 2007; van Ommeren and Gutierrez-i-Puigarnau, 2011). Furthermore, fixed-effects negative binomial models drop observations with zero outcomes. Thus, we do not consider negative binomial models as our main econometric models.

effect of commuting on sick-day absenteeism, with an elasticity of around 0.016. We investigate gender differences in this relationship, and find that it is statistically significant only for male workers.

The results presented in this paper may be of interest for firms and policy makers. For firms, sick-day absenteeism is costly, as it directly affects the firm's labor cost. Thus, firms should investigate to what extent reducing the commuting of their workers results in decreases in their costs as a consequence of a reduction in sick-day absenteeism. However, reducing the commuting of their workers may be probably achieved at the expense of higher compensation costs (Shapiro and Stiglitz, 1984) for the commuting of their workers (i.e., higher wages so that workers use their own cars rather than public transport, for instance), which may lead to an increase in the firm's labor cost. Given these two opposite effects on labor costs, firms could investigate the optimal compensation for the commuting of their workers that minimizes labor costs. From the point of view of policy-makers, improvements in infrastructure that allow for faster and/or shorter commutes may be beneficial for both firms and workers, which justifies national and local governments investing in the improvement of transportation networks and services.

One limitation of the paper is that, although we use panel data, we cannot talk about a causal link between commuting and sick-day absence, because we cannot control for time-varyiant unobserved factors that may be related to both commuting behavior and sick-day absence. Furthermore, despite that we control for health status and changes in job and residence location of workers, in an attempt to isolate the potential case of searching for work closer to home, because of poor health status, there may still be unobserved factors that affect these variables and sick-day absence. Finally, endogeneity due to measurement errors is present in longitudinal data models. For all these reasons, we cannot establish a causal link between commuting time and sickday absence, and more research on this topic is needed.

REFERENCES

Aguiar, M. & Hurst, E. (2007). Measuring trends in leisure: The allocation of time over five decades. *Quarterly Journal of Economics*, 122, 969-1007.

Allen, S. G. (1983). How much does absenteeism cost? Journal of Human Resources, 379-393.

Allison, P. D., & Waterman, R. P. (2002). Fixed-effects negative binomial regression models. Social Methodology, 32(1), 247-265.

Baker, C. C., & Pocock, S. J. (1982). Ethnic differences in certified sickness absence. Occupational and Environmental Medicine, 39(3), 277-282.

Barmby, T., Orme, C., & Treble, J. (1995). Worker absence histories: a panel data study. Labour Economics, 2(1), 53-65.

Barmby, T. A., Session, J., & Tremble, J. B. (1994). Absenteeism, efficiency wages and shirking. Scandinavian Journal of Economics, 96(4), 561-566.

Bianchi, S. M., Milkie, M. A., Sayer, L. C., & Robinson, J. P. (2000). Is anyone doing the housework? Trends in the gender division of household labor. Social Forces, 79(1), 191–228.

Billor, N., Hadi, A. S., & Velleman, P. F. (2000). BACON: blocked adaptative computationally efficient outlier nominators. Computation Statistics & Data Analysis, 34(3), 279–289.

Bonke, J. (2005). Paid work and unpaid work: Diary information versus questionnaire information. Social Indicators Research, 70(3), 349–368.

Bratberg, E., & Naz, G. (2009). Does paternity leave affect mothers' sickness absence? Bergen, Norway: University of Bergen Economics WP 06/09.

Campaña, J, C., Gimenez-Nadal, J. I., & Molina, J. A. (2016). Increasing the human capital of children in Latin American countries: The role of parent's time in childcare. The Journal of Development Studies, 53(6), 805-825.

Carta, F., & De Philippis, M. (2018). You've come a long way, baby. Husbands' commuting time and family labour supply. Regional Science and Urban Economics, 69, 25-37.

Dickerson, A., Hole, A. R., & Munford, L. A. (2014). The relationship between well-being and commuting revisited: Does the choice of methodology matter? Regional Science and Urban Economics, 49, 321-329.

Frey, B. S., & Stutzer, A. (2008). Stress that doesn’t pay: the commuting paradox. Scandinavian Journal of Economics, 110, 339–366.

Gershenson, S. (2013). Do summer time-use gaps vary by socioeconomic status? American Educational Research Journal, 50(6), 1219-1248.

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.

Gimenez-Nadal, J. I., & Molina, J. A. (2016). Commuting time and household responsibilities: Evidence using propensity score matching. Journal of Regional Science, 56(2), 332–359.

Gimenez-Nadal, J. I., Molina, J. A., & Velilla, J. (2018a). The commuting behavior of workers in the United States: Differences between the employed and the self-employed. Journal of Transport Geography, 66, 19-29.

Gimenez-Nadal, J. I., Molina, J. A., & Velilla, J. (2018). Spatial distribution of US employment in an urban efficiency wage setting. Journal of Regional Science, 58, 141-158.

Gimenez-Nadal, J. I., & Sevilla, A. (2011). The time-crunch paradox. Social Indicators Research 102, 181–196.

Gimenez-Nadal, J. I., & Sevilla, A. (2012). Trends in time allocation: A cross-country analysis. European Economic Review, 56(6), 1338–1359.

Goerke, L., & Lorenz, O. (2017). Commuting and sickness absence. IZA Discussion Paper No. 11183.

Gordon, P., Kumar, A., & Richardson, H. (1989). Gender differences in metropolitan travel behavior. Regional Studies, 23, 499–510.

Gottholmseder, G., Nowotny, K., Pruckner, G. J., & Theurl, E. (2009). Stress perception and commuting. Health Economics, 18, 559-576.

Greene, W. (2007). Fixed and random effects models for count data. Mimeo.

Grinza, E., & Rycx, F. (2018). The Impact of Sickness Absenteeism on Productivity: New Evidence from Belgian Matched Panel Data. IZA DP 11543.

Gutiérrez-i-Puigarnau, E., & van Ommeren, J. N. (2010). Labour supply and commuting. Journal of Urban Economics, 68(1), 82-89.

Hämmig, O., & Bauer, G. F. (2013). The social gradient in work and health: a cross-sectional study exploring the relationship between working conditions and health inequalities. BMC Public Health, 13(1), 1170.

Hanson, S., & Johnston, I. (1985). Gender differences in work trip lengths: implications and explanations. Urban Geography, 6, 193–219.

Hanson, S., & Pratt, G. (1995). Gender Work and Space. International Studies of Women and Place, London: Routledge.

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, 11, 834.

Jara-Díaz, S., & Rosales-Salas, J. (2015). Understanding time use: daily or weekly data? Transportation Research Part A: Police and Practice, 76, 38–57.

Juster, T., & Stafford, F. P. (1985). Time, goods, and well-being. Ann Arbor, MI: Institute for Social Research, University of Michigan.

Kaikkonen, R., Härkänen, T., Rahkonen, O., Gould, R., & Koskinen, S. (2015). Explaining educational differences in sickness absence: a population-based follow-up study. Scandinavian Journal of Work, Environment & Health, 41(1), 338-346.

Kirby, D.K., & LeSage, J. =. (2009). Changes in commuting to work times over the 1990 to 2000 period. Regional Science and Urban Economics, 39, 460-471.

Künn-Nelen, A. (2016). Does commuting affect health? Health economics, 25(8), 984-1004.

Leigh, J. P. (1991). Employee and job attributes as predictors of absenteeism in a national sample of workers: the importance of health and dangerous working conditions. Social Science & Medicine, 33(2), 127-137.

Löve, J., Hensing, G., Holmgren, K., & Torén, K. (2013). Explaining the social gradient in sickness absence: a study of a general working population in Sweden. BMC Public Health, 13(1), 545.

Ma, K. M., & Banister, D. (2006). Excess commuting: a critical review. Transport Reviews, 26, 749–767.

Markussen, S., Røed, K., Røgeberg, O. J., & Gaure, S. (2011). The anatomy of absenteeism. Journal of Health Economics, 30(2), 277-292.

Mastekaasa, A. (2000). Parenthood, gender and sickness absence. Social Science & Medicine, 50(12), 1827-1842.

Novaco, R. W., & Gonzalez, O. I (2009). Commuting and well-being. In: Amichai-Hamburger, Y. (Ed.), Technology and Psychological Well-Being. Cambridge University Press, Cambridge, U.K., pp. 174–205.

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(1), 1–8.

Van Ommeren, J. N., & Van der Straaten, J. W. (2008). The effect of search imperfections on commuting behavior: evidence from employed and self-employed workers. Regional Science and Urban Economics, 38, 127–147.

Piha, K., Laaksonen, M., Martikainen, P., Rahkonen, O., & Lahelma, E. (2009). Interrelationships between education, occupational class, income and sickness absence. European Journal of Public Health, 20(3), 276-280.

Piha, K., Laaksonen, M., Martikainen, P., Rahkonen, O., & Lahelma, E. (2012). Socio-economic and occupational determinants of work injury absence. European Journal of Public Health, 23(4), 693-698.

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.

Shapiro, C., & Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. American Economic Review, 74(3), 185–214.

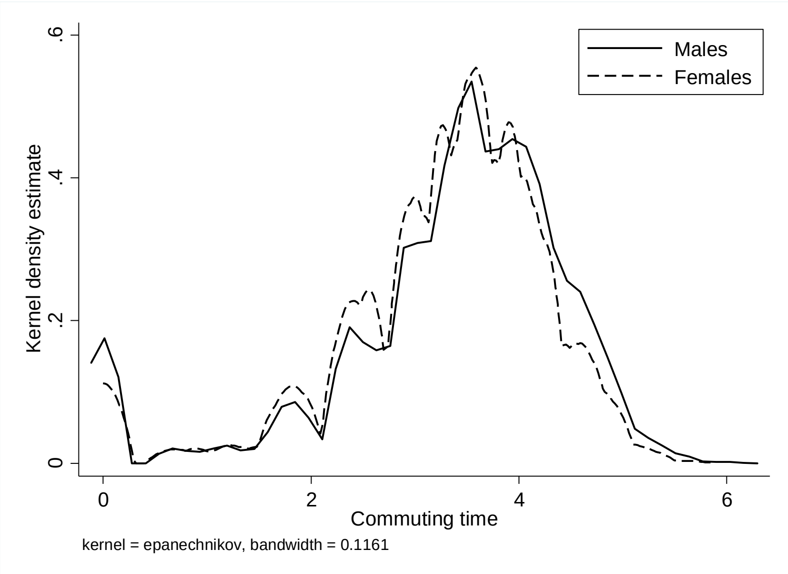
Simonsen, M., & Skipper, L. (2012). The family gap in wages: What wombmates reveal. Labour Economics, 19(1), 102-112.

Small, K., & Song, S. (1992). Wasteful commuting: a resolution. Journal of Political Economy, 100, 888–898.

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, 203–220.

Yee-Kan, M. (2008). Measuring housework participation: The gap between “stylised” questionnaire estimates and diary-based estimates. Social Indicators Research, 86(3), 381–400.

Figure 1. K-density of log-of-commuting times



Note: The sample (PSID 2011-2015) is restricted to workers who report positive hours of market work. Self-employed workers are excluded. Commuting time is measured in log-of-minutes per day.

Table 1. Summary Statistics, by gender

	Men		Women	
	Mean	S.D.	Mean	S.D.
Sick-day absence	0.778	2.704	0.744	2.366
Between variation		1.763		1.932
Within variation		2.061		1.612
Commuting time	43.196	39.076	37.532	32.650
Between variation		32.668		28.633
Within variation		22.870		16.940
Zero commuter	0.061	0.239	0.045	0.207
Age	41.791	11.590	42.515	11.334
Years of education	13.832	2.290	14.268	2.285
Hours worked per week	44.463	11.155	38.008	10.929
Annual salary	59614	114289	38810	31340
N. of children ≤ 6 years	0.397	0.732	0.402	0.717
N. of children 7-17 years	0.665	0.986	0.769	1.015
Live in couple	0.783	0.412	0.999	0.031
Being white	0.679	0.467	0.718	0.450
Total family income	94680	126764	105309	74552
Health: Excellent	0.212	0.409	0.174	0.379
Health: Very good	0.409	0.492	0.414	0.493
Health: Good	0.293	0.455	0.323	0.468
Health: Fair	0.074	0.262	0.077	0.267
Health: Poor	0.012	0.108	0.012	0.108
Moved residence	0.337	0.473	0.253	0.435
New job	0.076	0.264	0.054	0.226
Observations	8,649		6,245	
Individuals	3,306		2,704	

Note: The sample (PSID 2011-2015) is restricted to workers who report positive hours of market work. Self-employed workers are excluded. Commuting time is measured in minutes per day. Age is measured in years. Annual salary and Total family income are measured in dollars. Live in couple takes value 1 if individuals live with a spouse, or an unmarried partner, and 0 otherwise. Moved residence takes value 1 if the individual has moved during the year of the corresponding interview, 0 otherwise. New job takes value 1 if the individual has started a new job the year of the corresponding interview, 0 otherwise.

Table 2. Fixed Effects estimates

	Baseline model			Plus controls		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**General**	**Men**	**Women**	**General**	**Men**	**Women**
Log-commuting time	0.016**	0.017**	0.014	0.016**	0.018**	0.014
	(0.007)	(0.008)	(0.013)	(0.007)	(0.008)	(0.013)
Age	0.015	0.015	0.008	0.016	0.014	0.011
	(0.025)	(0.032)	(0.038)	(0.024)	(0.032)	(0.038)
Age squared	-0.009	0.006	-0.024	-0.007	0.008	-0.024
	(0.015)	(0.018)	(0.025)	(0.015)	(0.018)	(0.025)
Years of education	-0.006	-0.015	0.012	-0.005	-0.016	0.013
	(0.020)	(0.025)	(0.036)	(0.020)	(0.025)	(0.037)
Hours worked per week	0.005**	0.005	0.002	0.005**	0.005	0.002
	(0.002)	(0.003)	(0.004)	(0.002)	(0.003)	(0.004)
Hours worked per week sq.	-0.003	-0.003	-0.000	-0.003	-0.003	0.000
	(0.002)	(0.003)	(0.005)	(0.002)	(0.003)	(0.005)
Annual salary	0.000	0.000	0.000***	0.000	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annual salary squared	-0.000	-0.000	-0.000**	-0.000	-0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N. of children ≤ 6 years	-0.024	-0.014	-0.033	-0.025	-0.014	-0.033
	(0.017)	(0.021)	(0.030)	(0.017)	(0.020)	(0.030)
N. of children 7-17 years	-0.007	-0.001	-0.015	-0.006	0.000	-0.013
	(0.018)	(0.023)	(0.029)	(0.018)	(0.023)	(0.029)
Live in couple	-0.046	-0.047	-0.029	-0.047	-0.047	-0.043
	(0.037)	(0.038)	(0.447)	(0.037)	(0.038)	(0.441)
Being white	-0.207	-0.228	-	-0.201	-0.219	-
	(0.159)	(0.161)		(0.157)	(0.157)	
Log-total family income	0.173*	0.205*	-0.061	0.166*	0.198*	-0.107
	(0.089)	(0.107)	(0.378)	(0.088)	(0.106)	(0.393)
Log-total family income sq.	-0.687	-0.831	0.224	-0.663	-0.804	0.427
	(0.452)	(0.559)	(1.692)	(0.450)	(0.553)	(1.755)
Health: Very good	-	-	-	0.025	0.046*	-0.009
				(0.019)	(0.024)	(0.030)
Health: Good	-	-	-	0.023	0.049	-0.022
				(0.023)	(0.030)	(0.037)
Health: Fair	-	-	-	0.091**	0.126***	0.024
				(0.037)	(0.049)	(0.055)
Health: Poor	-	-	-	0.079	0.245	-0.207
				(0.113)	(0.154)	(0.140)
Moved residence	-	-	-	-0.007	-0.002	-0.010
				(0.017)	(0.021)	(0.028)
New job	-	-	-	-0.058**	-0.068**	-0.032
				(0.024)	(0.028)	(0.046)
Constant	-0.540	-0.378	0.065	-0.570	-0.361	0.198
	(1.186)	(1.428)	(2.709)	(1.179)	(1.411)	(2.781)
State F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Occupation F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14,894	8,649	6,245	14,894	8,649	6,245
Individuals	6,010	3,306	2,704	6,010	3,306	2,704
R-squared	0.012	0.015	0.021	0.013	0.019	0.023

Note: Robust standard errors in parentheses. The sample (PSID 2011-2015) is restricted to workers who report positive hours of market work. Self-employed workers are excluded. The dependent variable is the log of sick-day absences. Commuting time is measured in log-of-minutes per day. Age is measured in years. Annual salary and total family income are measured in dollars. Live in couple takes value 1 if individuals live with a spouse, or an unmarried partner, and 0 otherwise. Moved residence takes value 1 if the individual has moved during the year of the corresponding interview, 0 otherwise. New job takes value 1 if the individual has started a new job the year of the corresponding interview, 0 otherwise. Squared explanatory variables are defined as the square of the corresponding variable, divided by 100. *** significance at the 1%, ** significance at the 5%, * significance at the 10%.

Table 3. Robustness checks

VARIABLES	(1) General	(2) Men	(3) Women
A. Reduced sample			
Log-commuting time	0.015**	0.017**	0.012
	(0.007)	(0.009)	(0.013)
Observations	14,534	8,397	6,137
Individuals	5,963	3,278	2,685
B. Negative Binomial Fixed Effects			
Log-commuting time	0.062**	0.103***	-0.010
	(0.026)	(0.032)	(0.047)
Observations	5,823	3,644	2,179
Individuals	2,047	1,268	779
C. Commuter workers only			
Log-commuting time	0.020*	0.024*	0.016
	(0.010)	(0.013)	(0.017)
Observations	13,241	7,548	5,693
Individuals	5,270	2,833	2,437
D. Including self-employed workers			
Log-commuting time	0.014**	0.017**	0.007
	(0.006)	(0.007)	(0.010)
Observations	16,824	9,871	6,953
Individuals	6,571	3,632	2,939

Note: Robust standard errors in parentheses. The reduced sample (PSID 2011-2015) is restricted to workers who report positive hours of market work. Additional restrictions: A) The sample is restricted to individuals who report less than 15 sick-day absences, and less than 120 minutes of commuting per day. Self-employed workers are excluded. B) Self-employed workers are excluded. C) Zero-commuters are excluded. Self-employed workers are excluded. D) Self-employed workers are included. The dependent variable is the log of sick-day absence in models (A), (C), and (D); and the sick-day absences in model (B). Commuting time is measured in log-of-minutes per day. Additional estimates are shown in Table A1, in the Appendix. *** significance at the 1%, ** significance at the 5%, * significance at the 10%.

Appendix A

Table A1. Robustness checks, additional results

	A. Reduced sample			B. Negative Binomial Fixed Effects			C. Commuter workers only			D. Including self-employed workers		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	General	Men	Women	General	Men	Women	General	Men	Women	General	Men	Women
Log-commuting time	0.015**	0.017**	0.012	0.062**	0.103***	-0.010	0.020*	0.024*	0.016	0.014**	0.017**	0.007
	(0.007)	(0.009)	(0.013)	(0.026)	(0.032)	(0.047)	(0.010)	(0.013)	(0.017)	(0.006)	(0.007)	(0.010)
Age	0.028	0.024	0.027	0.023	-0.028	0.122**	0.021	0.017	0.017	0.028	0.027	0.031
	(0.024)	(0.031)	(0.037)	(0.031)	(0.039)	(0.057)	(0.026)	(0.034)	(0.040)	(0.022)	(0.028)	(0.034)
Age squared	-0.011	0.001	-0.023	-0.035	0.013	-0.129**	-0.008	0.003	-0.015	-0.001	0.014	-0.027
	(0.014)	(0.017)	(0.024)	(0.035)	(0.045)	(0.063)	(0.016)	(0.020)	(0.026)	(0.013)	(0.016)	(0.022)
Years of education	-0.018	-0.033	0.004	0.015	0.086***	-0.138***	-0.005	-0.007	-0.009	0.006	0.003	0.007
	(0.021)	(0.026)	(0.033)	(0.025)	(0.033)	(0.046)	(0.021)	(0.026)	(0.037)	(0.018)	(0.022)	(0.035)
Hours worked per week	0.005**	0.006**	0.001	0.034***	0.044**	0.011	0.004*	0.005	-0.001	0.006***	0.007**	0.004
	(0.002)	(0.003)	(0.004)	(0.013)	(0.018)	(0.021)	(0.003)	(0.003)	(0.004)	(0.002)	(0.003)	(0.003)
Hours worked per week sq.	-0.003	-0.004	0.001	-0.031**	-0.040**	-0.012	-0.003	-0.003	0.003	-0.004**	-0.005*	-0.001
	(0.002)	(0.003)	(0.005)	(0.013)	(0.018)	(0.025)	(0.003)	(0.003)	(0.005)	(0.002)	(0.002)	(0.004)
Annual salary	0.000	0.000	0.000***	0.000***	0.000***	0.000***	0.000	0.000	0.000***	0.000	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annual salary squared	-0.000	-0.000	-0.000*	-0.000***	-0.000**	-0.000***	-0.000	-0.000	-0.000***	-0.000	-0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N. of children ≤ 6 years	-0.023	-0.010	-0.036	-0.104**	-0.093	-0.073	-0.027	-0.022	-0.028	-0.022	-0.014	-0.035
	(0.017)	(0.020)	(0.028)	(0.051)	(0.065)	(0.090)	(0.018)	(0.022)	(0.032)	(0.016)	(0.019)	(0.027)
N. of children 7-17 years	-0.007	0.002	-0.016	0.012	0.024	0.042	-0.017	-0.013	-0.018	-0.001	0.013	-0.023
	(0.018)	(0.023)	(0.029)	(0.042)	(0.055)	(0.071)	(0.019)	(0.025)	(0.031)	(0.016)	(0.020)	(0.027)
Live in couple	-0.045	-0.047	-0.028	-0.219**	-0.133	-0.421	-0.059	-0.060	0.013	-0.054	-0.055	-0.022
	(0.037)	(0.038)	(0.450)	(0.109)	(0.117)	(0.902)	(0.040)	(0.041)	(0.451)	(0.034)	(0.035)	(0.428)
Being white	-0.185	-0.209	-	0.327***	0.131	0.909***	-0.036	-0.053	-	-0.279*	-0.288*	-
	(0.221)	(0.222)		(0.123)	(0.154)	(0.240)	(0.138)	(0.144)		(0.154)	(0.156)	
Log-total family income	0.161*	0.176*	-0.007	2.338***	2.919***	0.456	0.152*	0.188*	-0.050	0.115	0.115	0.089
	(0.090)	(0.103)	(0.386)	(0.780)	(0.960)	(2.132)	(0.088)	(0.107)	(0.412)	(0.096)	(0.108)	(0.290)
Log-total family income sq.	-0.666	-0.742	-0.031	-10.72***	-13.47***	-2.098	-0.627	-0.769	0.120	-0.449	-0.431	-0.454
	(0.457)	(0.542)	(1.733)	(3.597)	(4.513)	(9.425)	(0.457)	(0.577)	(1.836)	(0.458)	(0.525)	(1.307)
Health: Very good	0.017	0.027	-0.002	0.111	0.203**	-0.024	0.026	0.045*	-0.009	0.027	0.046**	-0.008
	(0.018)	(0.023)	(0.029)	(0.075)	(0.093)	(0.131)	(0.020)	(0.027)	(0.031)	(0.018)	(0.023)	(0.028)
Health: Good	0.021	0.043	-0.015	0.104	0.232**	0.002	0.019	0.054	-0.037	0.031	0.050*	-0.012
	(0.023)	(0.029)	(0.037)	(0.085)	(0.105)	(0.150)	(0.025)	(0.033)	(0.039)	(0.022)	(0.027)	(0.035)
Health: Fair	0.058*	0.094**	-0.005	0.276**	0.472***	0.083	0.086**	0.123**	0.018	0.107***	0.142***	0.038
	(0.035)	(0.046)	(0.054)	(0.120)	(0.152)	(0.204)	(0.039)	(0.052)	(0.058)	(0.034)	(0.044)	(0.054)

Health: Poor	-0.027	0.136	-0.263**	0.189	0.711**	-0.349	0.101	0.307*	-0.220	0.071	0.192	-0.160
	(0.092)	(0.121)	(0.132)	(0.235)	(0.299)	(0.396)	(0.124)	(0.172)	(0.150)	(0.098)	(0.131)	(0.130)
Moved residence	-0.007	-0.005	-0.005	-0.063	-0.081	-0.009	-0.009	-0.001	-0.012	-0.005	-0.001	-0.008
	(0.016)	(0.021)	(0.026)	(0.058)	(0.071)	(0.105)	(0.018)	(0.022)	(0.030)	(0.015)	(0.019)	(0.026)
New job	-0.043*	-0.062**	-0.003	-0.213**	-0.270**	-0.129	-0.063**	-0.067**	-0.043	-0.046**	-0.058**	-0.016
	(0.023)	(0.027)	(0.042)	(0.103)	(0.126)	(0.184)	(0.026)	(0.030)	(0.050)	(0.021)	(0.026)	(0.041)
Constant	-0.775	-0.291	-0.630	-15.51***	-19.27***	5.408	-0.300	-0.518	-0.042	-1.073	-1.037	-1.536
	(1.187)	(1.415)	(2.726)	(4.331)	(5.221)	(82.449)	(1.169)	(1.494)	(2.935)	(1.106)	(1.327)	(2.224)
State F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14,534	8,397	6,137	5,823	3,644	2,179	13,241	7,548	5,693	16,824	9,871	6,953
Individuals	5,963	3,278	2,685	2,047	1,268	779	5,270	2,833	2,437	6,571	3,632	2,939
R-squared	0.013	0.019	0.025	-	-	-	0.014	0.020	0.023	0.013	0.019	0.021

Note: Robust standard errors in parentheses. The sample (PSID 2011-2015) is restricted to workers who report positive hours of market work. Additional restrictions: A) The sample is restricted to individuals who report less than 15 sick-day absences, and less than 120 minutes of commuting per day. Self-employed workers are excluded. B) Self-employed workers are excluded. C) Zero-commuters are excluded. Self-employed workers are excluded. D) Self-employed workers are included. The dependent variable is the log of sick-day absences in models (A), (C), and (D); and the sick-day absences in model (B). Commuting time is measured in log-of-minutes per day. Age is measured in years. Annual salary and total family income are measured in dollars. Live in couple takes value 1 if individuals live with a spouse, or an unmarried partner, and 0 otherwise. Moved residence takes value 1 if the individual has moved during the year of the corresponding interview, 0 otherwise. New job takes value 1 if the individual has started a new job the year of the corresponding interview, 0 otherwise. Squared explanatory variables are defined as the square of the corresponding variable, divided by 100. *** significance at the 1%, ** significance at the 5%, * significance at the 10%.