The Impact of Cost Sharing for Mental Health Care on Mental Health Care Use and Labor Market Outcomes

Bastian Ravesteijn and Eli Schachar

November 7, 2018

Please click here for the most recent version.

Abstract

This paper investigates whether less generous insurance coverage for mental health care impacts economic productivity through reduced mental health treatment. We evaluate a health insurance reform in the Netherlands and use exogenous variation in the out-of-pocket price of treatment continuation, which depended on the exact date on which a patient began a prior treatment record before the reform was announced. We construct a dataset that links at the individual level administrative data on the mental health care claims records of all residents of the Netherlands to administrative data on employment. Our results indicate that the reduction in moral hazard under less generous insurance came at the expense of reductions in employment for certain subpopulations. Therefore, a well-targeted Pigouvian subsidy in the form of a lower out-of-pocket price for mental health care can be welfare-improving.

Ravesteijn: Harvard University and Paris Descartes University. Email: ravesteijn@hcp.med.harvard.edu. Schachar: Harvard University. Email: eschachar@fas.harvard.edu. We are extremely grateful to our advisors David Cutler, Eddy van Doorslaer, Tom McGuire, Claudia Goldin, and Larry Katz. We also thank Josh Angrist, Katherine Baicker, Aartjan Beekman, Amitabh Chandra, Chris Cronin, Gordon Dahl, Lisa Dixon, Rudy Douven, Liran Einav, Amy Finkelstein, Richard Frank, Titus Galama, Ed Glaeser, Sherry Glied, John Gruber, Ben Handel, Nathan Hendren, John Hsu, Haiden Huskamp, Richard Janssen, Bapu Jena, Patrick Jeurissen, Tim Layton, Tom Lee, Maarten Lindeboom, Robin McKnight, Michael McWilliams, Magne Mogstad, Joe Newhouse, Hessel Oosterbeek, Nick Papageorge, Mandy Pallais, Brian Perry, Harold Pincus, Erik Plug, Daniel Prinz, Thomas Rapp, Mark Shephard, Pietro Tebaldi, Bas van der Klaauw, Hans van Kippersluis, Dinand Webbink, and seminar participants at Harvard, the NBER Summer Institute, the AEA Annual Meeting, the ASHE Biennial Meeting, the IZA/SOLE Transatlantic Meeting, the University of Chicago, and Columbia University. This research was made possible by a data-use agreement between the authors and Statistics Netherlands (CBS), which has created de-identified datasets that are selectively available for use by academic researchers. Technical support was provided by Martin Broxterman, Ivo Gorissen, and Michael Vermaesen at CBS. The opinions expressed are those of the authors alone and do not represent the views of CBS. We gratefully acknowledge funding from The Dutch Ministry of Social Affairs and Employment, the Commonwealth Fund, the Pershing Square Fund for Research on the Foundations of Human Behavior, the Lab for Economic Applications and Policy Research, and the Department of Economics at Harvard University through the Warburg Prize.

1 Introduction

A key question in health economics is how to design health insurance contracts that minimize wasteful health care spending, while protecting individuals against the financial risk of negative health shocks. Increasing the amount patients pay out of pocket, known as cost sharing, could reduce the use of care that patients do not value highly, but which they nevertheless consume when it is covered by health insurance (Arrow, 1963). In particular, the demand for mental health care has been shown to be relatively elastic, suggesting that individuals don't value some of these services highly (Frank and McGuire, 2000). However, individuals potentially do not fully take into account that untreated mental illness may deteriorate human capital (Currie and Madrian, 1999). Moreover, they do not bear the full cost of loss of employment because of income insurance, which affects their incentives to invest in mental health. This implies that the reductions in wasteful spending when cost sharing for mental health care is increased may come at the cost of a loss in employment. Existing research, which has been hindered by a lack of (quasi-)experimental settings and adequate data, has offered limited evidence on the impact of cost sharing for mental health care on economic outcomes.

Our paper fills this void by exploiting a 2012 reform in the Netherlands to estimate the effect of higher cost sharing for mental health care on employment. This national reform created variation in patient cost sharing and led individuals at the margin to forgo mental health treatment. Our data consists of detailed mental health treatment records for all residents linked at the individual level to tax records containing the start and end dates of all employment contracts in the Netherlands. Differentiating by severity of mental illness and socioeconomic status, we show that greater patient cost sharing for mental health led to a reduction in months employed over the subsequent year for certain types of patients. Insurance design could take this externality into account by reducing patient cost sharing for these precisely defined types of patients.

We construct a new dataset to overcome a significant challenge to understanding the impact of cost sharing for mental care, namely the lack of comprehensive data that links mental health treatment status to labor market outcomes at the individual level. For each individual from the full Dutch population of 17 million for the period 2010 through 2012, we observe administrative data that contain detailed information on an individual's use of mental health care, including the start and end dates of treatment records, DSM-IV diagnosis codes and the type of treatment that was provided. We link these data to administrative data containing detailed information on the universe of employment contracts and social insurance claims, as well as demographic information and educational outcomes. In addition,

we observe measures of an individual's mental health for a subset of the population from a large 2012 survey and from separate, annual cross-sectional surveys from 2001 through 2015. These data provide the breadth and depth necessary to estimate reliably important indicators of how mental health care affects employment.

Our empirical strategy relies on the fact that the 2012 national reform in the Netherlands introduced variation in cost sharing that depended on choices the individual made before the reform was announced. To understand our identification strategy and the margin at which it operates, consider that the Dutch government sets the annual level of patient cost sharing, which affects the full, universally insured population. In 2011, the single health-care-wide annual deductible, which applied to total costs summed across all types of specialist health care, was €170. Amid concerns that spending on mental health care was rising too rapidly, the Dutch national government increased the out-of-pocket price for outpatient and inpatient specialist mental health care effective January 1, 2012. For outpatient care, adult patients were required to pay €100 out of pocket to open a treatment record and an additional €100 if they received a total of more than 100 minutes of care during that treatment record. For inpatient care, adult patients were required to pay a monthly copay of €145. A mental health treatment record worked as follows. The physician opened a treatment record when the patient started treatment, and a single treatment record covered all of the patient's care for up to 365 days after the date on which the record was opened. However, all treatment records automatically closed after 365 days; to continue treatment, a patient then had to open a new treatment record. Consequently, patients were affected by the post-reform out-of-pocket prices if and only if their treatment record was opened in 2012.

To estimate the impact of higher cost sharing on employment, we exploit this variation created by the combination of the cost-sharing reform and the timing of when treatment records were automatically closed. In particular, we compare individuals who opened treatment records in late 2010 with those who opened treatment records in early 2011, conditional on a smooth function of the date of opening. After 365 days, the treatment records for both the late 2010 group and the early 2011 group were automatically closed, and the individuals in both of those groups then faced the decision of whether or not to continue treatment by opening a new treatment record. However, the individuals in these two groups faced different prices for opening a new treatment record: those in the late 2010 group, who made their continuation decision when their record closed in late 2011, faced the pre-reform out-of-pocket price; those in the early 2011 group, who made their continuation decision when their record closed in early 2012, faced the post-reform out-of-pocket prices. We show that the late 2010 and early 2011 groups are similar on observable characteristics. Moreover, neither group could have timed the opening of their initial treatment record in

anticipation of the reform, because the reform was not announced until July 2011. We obtain first-stage estimates of the effect on treatment continuation and reduced-form estimates of the effect of the cost sharing reform on employment. The Wald estimate, the ratio of the reduced-form over the first-stage estimate, can be interpreted as a lower bound of the local average treatment effect of treatment continuation on employment. The validity of these results relies on the assumption that the conditional expectations of the potential outcome (employment) with respect to start date of the initial treatment is smooth through the January 1, 2011 threshold. The estimates can be interpreted as local average treatment effects of effect of mental health care treatment on employment. This is precisely the margin that is relevant for policy-makers, because it applies to individuals whose decision to use mental health care can be influenced by cost-sharing policies.

Our results indicate that greater cost sharing leads to loss of employment. The reduced-form effect of a €200 increase in the annual mental health care-specific out-of-pocket maximum among individuals who had initiated a treatment record in 2010 or 2011 is an estimated average employment loss of 3.7 days in the next year. Exploring the mechanism of this effect, we find that treatment continuation drops by nearly four percentage points because of the increase in cost sharing, relative to a baseline treatment continuation rate of 57%. Assuming that the effect on employment is driven by the individuals who decided not to continue treatment because of the additional cost sharing, the estimated local average treatment effect of treatment discontinuation is an employment loss of 14 weeks. These results suggest that mental health, like education and physical health, should be viewed as an important part of human capital which can depreciate and be repaired through mental health investment in the form of mental health care.

This paper builds on our prior work investigating the impact of the Dutch cost-sharing reform on mental health care use. Ravesteijn et al. (2017) used a difference-in-differences estimation strategy to show that the use of regular mental health care decreased abruptly and persistently by 13.4%, while use increased for involuntary commitment by 96.8% and for acute crisis mental health care by 25.1%, after the Dutch cost-sharing reform was implemented. Relative to that paper, this paper offers three significant innovations. First, this paper uses a dataset that not only contains the treatment records for the full population, but also links treatment records within an individual across time. This allows us to investigate treatment continuation patterns among individuals who had been in mental health care prior to the reform. Second, this paper links the data on mental health care use to employment-level data, so we can investigate offset effects outside of the mental health care sector. Finally, this paper uses an empirical strategy that is robust to strategic behavior, and thus, allows us to estimate the causal impact of higher patient cost sharing for mental

health care on labor market outcomes.

This paper also contributes to the existing literature on optimal cost sharing. Moral hazard refers to the greater quantities of health care that insured individuals use in excess of the health care costs that they would be willing to insure themselves against if the insurer could perfectly monitor health (Aron-Dine et al., 2013; Cutler and Zeckhauser, 2000; Manning et al., 1987). The classical moral hazard model takes the slope of the demand curve as a sufficient statistic to compare the welfare consequences of different levels of insurance coverage (Pauly, 1968; Zeckhauser, 1970). The reasoning is that a stronger demand reduction in response to a price increase indicates that patients value those services less. However, more recent research highlights two reasons for why the price elasticity by itself does not determine the socially optimal level of cost sharing. First, cost sharing may lead to offset effects elsewhere in the health care system. In the presence of multiple health repair technologies, optimal cost sharing must take cross-price elasticities into account because an increase in cost sharing for one health service may induce individuals to switch to other, possibly costlier, health services (Chandra et al., 2010, 2014; Finkelstein et al., 2012; Goldman and Philipson, 2007). Similarly, an increase in cost sharing could cause individuals to cut back on cost-effective preventive care and instead use expensive curative care once their health has deteriorated (Aron-Dine et al., 2015), which could lead to suboptimal intertemporal patterns of health care use. Second, individuals might not be aware of their need for health treatment or the returns of such treatment (Baicker et al., 2015; Pauly and Blavin, 2008), and this lack of information could lead individuals to consume levels of health care that do not maximize their utility. These theoretical predictions are confirmed by empirical findings that suggest that, in response to higher cost sharing, patients reduce their use of not only low-value but also high-value care (Brot-Goldberg et al., 2017; Newhouse and Rand Corporation Insurance Experiment Group, 1993). Our paper adds to the literature on offset effects, looking beyond the health care sector by taking into account that lack of treatment could lead to diminished economic productivity. In particular, our paper provides evidence that socially optimal health insurance should account for labor market externalities by lowering cost sharing for certain patients.

The remainder of the paper is organized as follows. Section 2 describes our data and documents that mental illness is strongly associated with worse labor market outcomes. Section 3 explains the reform and the empirical strategy we use to estimate the causal impact of higher patient cost sharing on employment. Section 4 presents our results, and the last section concludes.

2 Data on mental health, treatment, and employment

We constructed a dataset that allows us to investigate the relationship between mental health (treatment), cost sharing, and labor market outcomes.

2.1 Data sources and base sample

We linked the specialist mental health care treatment records to 15 additional data sources. We limit the sample to adults of working age (between 18 and 64 years old on January 1, 2011) with uninterrupted registration in the municipal base administration file between January 1, 2010 and December 31, 2012. This forms our base sample, to which we linked information on: all specialist mental health treatment records from DIS; K-10 and MHI-5 mental health scores from cross-sectional health surveys; monthly indicators of any income received from employment in that month for all months between 2010 and 2013; employment records containing information on hourly wage, an indicator of temporary employment contracts as opposed to permanent contracts, a wage cut due to prolonged illness, and part-time employment ($\leq$ 24 hours a week); annual information on the primary source of income, personal and household income, and household wealth of each Dutch resident from the Dutch tax office; information on DI claims and the diagnosis that was the main cause of disability between 2011 and 2013 from the Dutch social security administration; gender, year and month of birth, the country of birth of the parents and the individual him-or herself which we used to construct a measure of ethnicity, and the composition of each household; highest completed degree according to the ISCED 2011 classification of 5,295,690 adults; year and cause of death for 2014 and 2015 from the cause-of-death registry (DO); health care claims amounts by type of health care in 2011 from the national health insurance database institutionalization in nursing, disability, and mental health institutions from the long-term care registry. Appendix A describes the data sources in detail and explains how we constructed the variables.

2.2 Mental Health: Treatment Records and Kessler-10 scores

Our main source of information on mental health treatment is the full DBC Informatie Systeem (DIS), which contains all specialist mental health care treatment records in the Netherlands. Specialist mental health care is an essential benefit in all health insurance plans that are offered in the Netherlands, and health insurance is mandatory. As a rule, not more than one treatment record can be opened at any time, and treatment records are automatically closed after 365 days. The treatment records formed the basis of payments

to the provider by the health insurance company, which occurred after a treatment record had been closed. Treatment continuation beyond 365 days requires that the patient and the provider agree to open a new treatment record.

Table I shows descriptive statistics for all treatment records (not individuals) that were open at any time in 2011 of adults in our baseline sample. In total, 691,312 out of 10,172,176 adults of working age (6.80%) received specialist mental health treatment in 2011. These records do not contain information on long-term mental health care use, treatment provided by primary care physicians (PCPs) and by primary mental health care providers, or the type of prescribed pharmaceuticals. A treatment record is classified as "initial" if no prior treatment record with the same primary diagnosis code at the same provider exists. A "continued" treatment refers to records that follow a previous record at the same provider organization with the same primary diagnosis. A record is determined "involuntary commitment" if within the first month after the start of the record a BOPZ commitment procedure was initiated. In 2011, 622,474 initial regular treatment were open, versus 514,842 continued treatment records and 1,530 involuntary commitment records.

Panel A of Table I shows that, on average, patients at the beginning of an initial treatment record have a similar Global Assessment of Functioning (GAF, see Jones et al., 1995) scores as those beginning a continued treatment record. These scores range between 1 ("persistent danger of severely hurting self or others OR persistent inability to maintain minimal personal hygiene OR serious suicidal act with clear expectation of death") and 10 ("superior functioning in a wide range of activities, life's problems never seem to get out of hand, is sought out by others because of his or her many positive qualities; no symptoms"). GAF scores of committed individuals are approximately 4 ("some impairment in reality testing or communication OR major impairment in several areas, such as work or school, family relations, judgment, thinking, or mood") on average.

[Table I about here.]

GAF is assessed and recorded by the therapist at the beginning and at the end of each treatment record and on average improve over the course of a treatment record. The average increase is greater for initial regular treatment and involuntary commitment than for continued regular treatment. These changes in GAF should not be interpreted as the effect of treatment, because we do not observe what would have happened to the patient in absence of treatment. Therefore we cannot rule out that improvement in GAF would have occurred without treatment, or that GAF scores would have deteriorated in the absence of treatment.

Those in continued regular treatment had the longest average treatment duration, measured as the number of days between the start and end of a treatment record. Those in involuntary commitment had more treatment minutes, more visits (days with at least one recorded treatment activity), a higher probability of at least one inpatient day, more treatment days, and were more likely to have medication prescribed to them, inferred from the fact that treatment minutes were recorded for prescribing any kind of medication. Medication prescription was greater in continued than initial treatment records, and medication could also be prescribed by a PCP outside of the specialist mental health care system.

Panel B of Table I shows the distribution of twelve DSM-IV-based primary major diagnostic classes, which we constructed using 689 observed primary diagnosis codes. Bipolar disorder, personality disorder, and psychotic disorder were more common as primary diagnoses among continued regular and involuntary commitment treatment records compared with initial treatment records. Tables B.I and B.II in appendix B.1 show the information in Table I broken down by diagnosis. It shows that psychotic disorder, and to a lesser extent personality disorder and substance use, are associated with low GAF scores, more treatment minutes, and more medication prescription compared with the other nine major diagnostic classes. Adjustment disorders and other conditions that may be a focus of clinical attention had relatively high GAF scores and fewer treatment minutes on average.

Out of the 82,386 records with the major diagnosis class of disorder first diagnosed in childhood, 41,917 were recorded as attention deficit disorder, while 32,120 were recorded as pervasive developmental disorders such as autistic disorder and Asperger's disorder. Of the 78,934 miscellaneous disorders, 30,129 were somatoform disorders, 16,240 were eating disorders, 13,364 were impulse control disorders, and 5,204 were sexual and gender identity disorders. Of the 95,932 records classified as other conditions that may be a focus of clinical attention, 59,025 were classified as relationship problems, and 15,393 as identity problems.

Using the GEMON cross-sectional survey in 2012, we linked Kessler-10 mental health scores for 220,627 individuals in our baseline sample ranging between 10 (poor mental health) and 50 (excellent mental health), to the mental health care treatment record file. Panel A of Figure 1 shows that the distribution of K-10 scores is left-skewed and Panel B shows that around 30% of individuals with K-10 scores below 25 received specialist mental health treatment in 2011, whereas this was 1.05% among individuals with a maximum K-10 score of 50. A similar pattern can be found for scores based on the five-item Mental Health Inventory (MHI-5) questionnaire instead of the K-10 scores, which were recorded by the cross-sectional POLS (2001-2009) and GECON (2010-2015) surveys. In the years prior to and after 2011, mental health scores of individuals who were in treatment in 2011 were on average lower

than those who were not in treatment in 2011, and displayed a downward trend up to 2011, and an upward trend in the years after 2011. Appendix B.2 provides the empirical evidence for these claims about the relationship between MHI-5 scores and treatment status.

[Figure 1 about here.]

2.3 Associations between mental health (care use) and economic outcomes

We document some associations that motivate the central empirical question of this paper: can mental health care improve labor market outcomes? Column 1 of Table II shows the mean personal and household income and mean household wealth measured in 2011, for all adults of working age with a K-10 score of 43 and higher (72% of all individuals for whom we observe a K-10 score). In addition, it shows that the mean hourly wage among employed individuals in January 2011 with K-10 scores $\geq$ 43 is €19.52. Column 2 shows the differences between individuals with K-10 scores $<$ 43 and those with K-10 scores $\geq$ 43 for these variables, conditional on age and gender. Column 3 shows the means of the same variables for all individuals in our baseline sample who were not in specialist mental health treatment in 2011, and column 4 shows the conditional differences between those who were in treatment and those who were not. Individuals in poor mental health and/or in treatment have on average lower personal and household income, lower household wealth, and lower hourly wages. This raises the question whether mental health treatment could improve economic outcomes.

[Table II about here.]

Table III shows that individuals with poor mental health and/or in mental health treatment were less likely to be employed in 2012, more likely not to be employed in 2015 if they were employed in 2012, and less likely to be employed in 2015 if they were not employed in 2012. In addition, they are more likely to have a temporary instead of a permanent employment contract, more likely to have been absent from work because of illness, and are more likely to be part-time employment. We point out that the subsample of employed individuals in GEMON with recorded K-10 scores appears not to be a random subsample of the sample of all individuals who are employed. For example, the proportion of all employment contracts that is temporary is 26.5%, while for the subsample with nonmissing K-10 scores this is only 21.5%.

[Table III about here.]

Table IV shows that individuals with poor mental health and/or in treatment are less likely to work, more likely to receive unemployment insurance, disability insurance, and basic welfare. In appendix B.3 we show that poor mental health and mental health treatment are concentrated among individuals with lower personal income, household income, and wealth.

[Table IV about here.]

2.4 Associations between mental health (treatment) and demographic characteristics, education, health care claims and long-term care

Given that no previous studies have documented used administrative data to document the associations between mental health (treatment) and demographic characteristics, we briefly describe some associations in our data, a more detailed description can be found in appendices and B.5. K-10 scores are on average higher for women than for men, are low during adolescence, increase up to the age of 45, then slightly decrease up to the age of 55, after which they strongly increase and peak at the age of 65, and then they gradually decrease. Mental health care use peaks around the ages of 10, 35, and 85, bottoms out between ages 65 and 70, is greater for women than for men, and greater for ethnic minorities. Women and ethnic minorities are also in worse mental health on average. Both mental health and treatment exhibit a strong gradient in education, with high educational attainment being associated with higher mental health scores and lower mental health care use. Cohabiting couples with and without children have higher mental health scores and lower use of mental health care compared with single parents, single adults, and residents of institutional homes.

Conditional on age and gender, both poor mental health and mental health care use are associated with a higher likelihood of death for each of the following causes: suicide, poisoning (accidental and intent unknown), organic mental disorders such as dementia, substance use disorders (not classified as poisoning), other mental disorders, and deaths by any other cause. While we lack power in the sample of individuals with nonmissing K-10 scores, mental health treatment is also associated with a higher probability of death by homicide.

Conditional on age and gender, poor mental health and mental health care use are associated with higher hospital, pharmaceutical, primary care physician, and other health care claims. Poor mental health and mental health treatment are associated with both higher residence in the three type of long-term care institutions (nursing, disability, and long-term care) and inflow in those institutions as well as a higher probability of long-term care admission. In the online-only appendix we break down all descriptive analyses by major diagnosis category.

2.5 Sample definition and missing values

Statistics Netherlands only makes available data for individuals who at some point resided in the Netherlands according to the Dutch Municipal Basic Administration (Dutch acronym GBA). We restrict our sample to individuals with uninterrupted registration in GBA between 2010 and 2012. Section A.14 C shows the number of observations for each additional step in the data linkage process: the number of observations in each data source; nonmissing values for each variable; nonmissing values for individuals who met the residency requirement; and nonmissing values for individuals who additionally met the age restriction.

3 Empirical Strategy

Our empirical strategy relies on variation in the out-of-pocket price of mental health care that was induced by a reform to patient cost sharing in the Netherlands in 2012. First, we explain the patient cost sharing reform in detail, which is necessary to understand the margins on which we expect the reform to affect individual and patient behavior. Second, we describe the identification method we employ to obtain from this national reform exogenous variation that allows us to identify causal effects.

3.1 The 2012 Patient Cost Sharing Reform

In July 2011, the Dutch national government announced that it would increase in 2012 the out-of-pocket price patients face for mental health care, including addiction treatment. The Dutch Minister of Health believed that a higher out-of-pocket price would discourage patients from using low-value behavioral health care and thus combat rising behavioral health care costs (Dutch Minister of Health, Welfare, and Sports, 2011), which since 2000 had grown at a rate two times higher than the rate of all other health care costs (see Figure 2).

[Figure 2 about here.]

Effective beginning January 1, 2012, patient cost sharing for mental health care increased in several ways. First, mental health patients aged 18 years or older were required to pay an annual copay of €100 for treatment lasting up to 100 minutes, and an additional €100 for treatment lasting longer, and a monthly copay for inpatient care of €145 euro. In addition, the health care-wide deductible was raised from €170 to €220 euro. As a result, patients aged 18 or older who required more than 100 minutes of outpatient mental health treatment and did not face other health expenditures went from facing an out-of-pocket cost of €170 in 2011 to an out-of-pocket cost of €420 when opening a treatment record in 2012.

Although the reform was implemented nationwide, certain patients were not affected by the reform, a key to our identification strategy. First, the out-of-pocket prices of primary care, medication, and non-mental health care were not affected by the reform, and if these services can also treat mental health the reform introduced an incentive to substitute to these services. Crisis care and involuntary commitment were not subject to cost sharing, and individuals who had started a treatment record in 2011 could continue treatment through 2012 without facing the new cost-sharing regime for the full duration of that treatment record, 365 days after its start date. Only after that grand-fathered treatment record was ended did they face the decision to open a new treatment record, subject to the new cost-sharing regime.

Simultaneously to the introduction of the copay, treatment of adjustment disorders was taken out of the set of essential health benefits and therefore no longer covered by all health insurance plan. In addition, it was announced that this would also happen for the DSM-IV main diagnostic class called "other conditions that may be a focus of clinical attention", which includes relationship and identity problems.

Figure 3 shows a sharp decrease in the number of treatment records by adults after January 1, 2012 while this number remained stable for youths, who were not affected by the cost sharing reform. In a previous paper, we report that the number of treatment record opening decreased by 13.4% (s.e. 1.3) after the introduction of the copay, and that this result is robust against controlling for youths who had not been subjected to the copay.

[Figure 3 about here.]

After the 2012 parliamentary elections in 2012, the new government decided to abolish the mental health care-specific copay on January 1, 2013. Simultaneously, the annual health care-wide deductible was increased from €220 in 2012 to €350 in 2013.

3.2 Identification

3.2.1 LATE as the policy-relevant marginal treatment effect

In this section we explain that the policy-relevant marginal treatment effect in our setting is the local average treatment effect (LATE) for individuals who are at the margin, and who only seek care in absence of (higher) cost sharing, and not the average treatment effect of mental health treatment on the full population. The LATE tells us whether treatment has an effect on employment for individuals who are at the margin of deciding whether to continue treatment and who change their behavior in response to a key policy lever, namely a modest increase in patient cost sharing at a comparatively low level of cost sharing. We then argue

that our identification strategy produces estimates that have precisely this interpretation. Our estimates are informative about the effect of treatment among individuals who adjust their behavior in response to an out-of-pocket price increase at the margin that was affected by the Dutch cost sharing reform.

A strategy that would provide us with a policy-relevant effect estimate is when a copay for mental health care (rather than mental health treatment itself) would be randomly assigned among the population. One could then study the impact of mental health treatment among compliers with the copay as an instrumental variable for mental health treatment: those who seek mental health treatment in the absence of the copay, but who forgo treatment if subject to the copay. Given that, under random assignment, the potential outcomes of the groups with and without the copay are equal in expectation, the difference in the probability of seeking treatment between both those with and without the copay could be interpreted as the first-stage (FS) effect of the copay on mental health care use. In addition, differences in employment can be interpreted as the reduced-form (RF) effect of the copay on employment. Under the exclusion restriction that the effect of the copay on employment only runs through mental health care use, and not through increased employment to pay for the copay, we could divide the reduced-form estimate by the proportion of individuals who changed their treatment-seeking behavior because of the effect, which is the first-stage estimate, to obtain the Wald estimator of the local average treatment effect of mental health treatment on employment.

If the exclusion restriction is violated because some individuals worked more to pay for the copay, this would lead to an upward bias of the Wald estimator because of the direct effect of the instrument on employment. If the local average treatment effect of mental health treatment on employment is positive because both the FS and RF coefficients are negative, the Wald estimator is biased towards zero. The local average treatment effect estimate would not tell us about the treatment effect for always-takers, who seek treatment regardless of the copay, and never-takers, who don't seek treatment even if it is not subject to the copay. Always-takers and never-takers are of less policy relevance because there is less worry of moral hazard, and their behavior is less likely to be influenced by policy. This strategy relies on the monotonicity assumption which rules out defiers, those who only seek treatment under the copay. This assumption would for example be violated if the copay was randomized by regional clusters rather than at the individual level, and that in regions with the copay waitlists would become much shorter such that individuals who would otherwise forgo care because of the long waiting time would now seek care.

3.2.2 Limitations of a before-after comparison

Given that we know of no setting with adequate data and random assignment of a copay for mental health, we looked for a setting in which the assignment of a copay was independent of potential outcomes. The reform in the Netherlands introduced a copay on January 1, 2012. However, it is not very informative to compare employment outcomes of the full population immediately before and after the reform for three reasons. First, the reform is not expected to immediately affect employment because we theoretically expect the effect of mental health care on employment to run through mental health, which is a stock variable that gradually adjusts over time. In the Netherlands, with its strict labor laws, those whose health deteriorates but who do not voluntarily leave their job are usually placed on sick leave which means that we do not expect an immediate effect on employment. As we have seen, the majority of employees have a permanent employment contract which can only be cancelled by the employer when a substantial severance payment is made: many workers on a temporary contract are entitled to sick pay for the full duration of the contract. Second, the introduction of the copay could be correlated with unobserved factors that vary over time, such that the potential outcomes before and after the reform are not the same and therefore the estimator of the effect of the copay is biased. Third, the reform is expected only to affect a very small portion of the population. As we have seen, only 6.80% of the Dutch adults of working age was using mental health care in 2011, and the proportion of adults who change their behavior because of the copay is expected to be much smaller than that. Therefore, even though we observe the treatment status and employment of the full Dutch population, lack of precision would prevent us from identifying variation in employment induced by variation in treatment-seeking behavior using this identification strategy.

We could improve the precision of our estimates by restricting the sample to individuals who are in poor mental health (as predicted by mental health scores or prior health care use), and therefore in expectation have a higher chance of seeking treatment. However, those in poor mental health in 2012 are likely to also be in need of treatment in 2011, and they could have anticipated the reform by starting treatment before the introduction of the copay on January 1, 2012, which would have allowed them to continue treatment for the full duration of that treatment record without being exposed to the reform. Depending on the magnitude of these anticipation effects, our resulting first stage estimate could be very small (because many individuals potentially got around paying the fixed access fee) and our Wald estimator would suffer from the weak instruments problem.

3.2.3 Variation in cost sharing for treatment continuation by the start date of a prior treatment record

If we could find comparable groups who randomly varied in their likelihood of facing the decision to open a treatment record either before or after the reform (but not both) for reasons unrelated to the copay or their potential outcomes, and follow their employment outcomes over time, this would get around the problems of delayed response and lack of precision that were described in the previous paragraph. We could then control for time trends in the exact date at which individuals became likely to open a treatment record before and after the reform, such that our estimators are based on a discontinuity in treatment uptake and employment outcomes right at the threshold of January 1, 2012, rather than reflecting secular employment trends over time that were correlated with timing of the decision to open a treatment record.

Our identification strategy does just that. We limit the sample to individuals who had opened a treatment record between August 1, 2010 and July 1, 2011. We will refer to this treatment record as the "prior record", although many of these individuals had a treatment record before opening this prior record. We then evaluate whether these patients had a treatment record that was open for at least 30 days during the six months beginning 365 days after the prior record was opened. We will refer to a record that was open in these six months as a follow-up record. Given that the maximum duration of a treatment record was 365 days, treatment in this follow-up period indicates that a new treatment record was opened.

Figure 4 illustrates our empirical setup. Patient A opens his prior record in November 2010, and after the maximum duration of that record in November 2011, he is not subject to the copay when opening a follow-up treatment record. In contrast, patient B opens his prior record in February 2011, and during the follow-up period (months 13 to 18 counting from the month in which the prior record was opened) was subject to the copay. We follow up both patients from the date 365 days after opening the prior record until six months after that date and consider him to be in follow-up treatment if he has an open record for at least 30 days during that six-month period.

[Figure 4 about here.]

3.2.4 Estimation equations

We estimate equation FS, where t is a dummy variable indicating whether someone was in continued treatment for at least 30 days in months 13-18, r is 1 if the prior record was opened in 2011 and 0 if it was opened in 2010, and $\mathbf{x}$ is a vector of age dummies; a running

variable indicating the day of opening of the prior treatment record counting from the first observed date; and that variable squared. β_2 refers the change in the probability of treatment continuation at the threshold.

$$t_i = \beta_1 + \beta_2 r_i + \boldsymbol{\beta}_3^{'} \mathbf{x}_i + \varepsilon_i \tag{FS}$$

As we explained, therapists and patients are formally not allowed to close a treatment record of less than 365 days if treatment will be continued in a newly opened record. However, we cannot rule out that some patients and providers colluded to strategically by closing a record and then open a new one to get around the fixed access fee. The reform was announced on July 26, 2011. We will investigate empirically whether patients and/or therapists have anticipated the reform by closing their prior treatment record and simultaneously opening a follow-up record. If this is the case, our estimates put a lower bound on the effect of the copay on treatment continuation.

We measure employment in the twelve months beginning 365 days after the prior record was opened. Equation RF is identical to equation FS with the exception of the outcome variable, which is now the proportion of months out of the twelve follow-up months in which the individual received income from employment. γ_2 refers to the decrease in the proportion of months employed at the threshold.

$$e_i = \gamma_1 + \gamma_2 r_i + \boldsymbol{\gamma}_3^{'} \mathbf{x}_i + \eta_i \tag{RF}$$

3.2.5 The Wald estimator

The ratio of γ_2 over β_2 gives us the Wald estimator. This could be viewed as a fuzzy regression discontinuity estimator conditional on the opening date of the prior treatment record (Hahn et al., 2001). The estimates can be interpreted as a local average treatment effect of treatment continuation on employment. The group of compliers for whom we estimate a local average treatment effect are those who continue receiving care in absence of the copay, but who forgo care if a copay is required.

Given that the reform was not announced until July 2011, patients and therapists could not have anticipated the reform by strategically setting the start date of the prior treatment record because that record had already been opened by the time the reform was announced. It is therefore unlikely that strategic behavior led to violation of the key identifying assumption, which is that the conditional expectations of the potential outcome (employment) with respect to start date of the prior treatment is smooth through the R =

0 threshold (Clark and Royer, 2013).

The group of compliers does not include individuals who were able to circumvent the copay and are therefore always-takers, for example by ending their prior treatment record early in order to start a continued record in late 2011. These individuals would continue treatment regardless of whether their prior treatment record had been opened in 2010 or 2011. Although this strategic behavior influences the composition of the group of compliers, it does not bias the estimator of the LATE for those compliers.

Given that the reform happened right at New Year and during the Dutch Christmas holiday, we investigate whether treatment records opened in the days around the cut-off date show abnormalities. To address potential concerns that there is non-random sorting around the reform date which may drive our results, we use a donut RD (Barreca et al., 2011) approach by removing records that were opened during the Dutch Christmas holidays.

3.2.6 Placebo test using matches

We create a sample of matches who had no mental health care spending in 2009 through 2014 and appeared in the baseline dataset. We performed exact 1-to-1 matching with replacement for our full analysis sample, matching on gender, age measured on January 1, 2010, and employment status in 2009. This matched sample was repeatedly observed over the follow up period, with each match being observed as many times as there were months in the analysis sample. We ran our analyses on this sample as a placebo test to investigate whether secular non-smooth trends through the threshold could explain any effect that we found in our main analyses. Unfortunately, we do not observe treatment records before 2010 or after 2012 which we could have used to perform placebo analyses on prior or later years.

3.2.7 Strengths and limitations

Our empirical strategy addresses several identification issues. First, in absence of a randomized controlled experiment, we evaluate a natural experiment that relied on real-world variation in cost sharing for mental health care. Second, rather than estimating average effects for the full or the treated population, we estimate a LATE of treatment continuation for people who are at the margin of deciding whether to seek treatment, and whose behavior can be influenced by (cost sharing) policies. Third, we limit the sample to individuals who are in substantial need of treatment (continuation), which improves the precision of our first-stage, reduced-form, and Wald estimators. Fourth, our analysis is not restricted to outcomes measured immediately before or after the reform, because by using variation in the timing of the decision to open a treatment record we are not hindered by the limitation

of a before-after comparison where all individuals are both in the "before" and in the "after" group. We can therefore follow outcomes of the 2010 and 2011 groups over a longer period of time. Fifth, even if the potential outcomes of the 2010 and 2011 groups are different, which we can partially investigate by testing for balance of observable characteristics, our identification strategy is still valid under the assumption that any variation in potential outcomes varies smoothly along the running variable (the opening date of the prior treatment record).

Several things are important to keep in mind when interpreting our results. First, because individuals in the 2011 group could still get around the copay, the first stage cannot be interpreted as an average treatment effect of the copay on treatment continuation for those who are in treatment. Instead, it places a lower bound on that effect. Second, while we are able to investigate heterogeneity of treatment effects by interacting r with GAF-scores or measure of socioeconomic status, our setting does not allow us to compare the effectiveness of different types of treatment (e.g. cognitive behavioral therapy versus psychotherapy versus pharmacotherapy) because the reform introduced variation in the costs for all specialist mental health care at once. Third, for individuals who placed a value of at least €200 on treatment of more than 100 minutes, the reform had no further effect on their behavior. Therefore, the reform can only be used to estimate treatment effects at the discrete "extensive" margin of opening a record, and not to answer a separate policy-relevant question, which is what the optimal treatment duration is after someone has elected to open a treatment record. Moral hazard at the intensive margin could mean that individuals seek treatment for more minutes than for which they would contract health insurance under perfect information, but our setting unfortunately does not allow us to investigate this further. Fourth, our LATE only applies to individuals who respond to cost sharing at the margin of the Dutch reform. They are uninformative about the potential tradeoff between moral hazard, risk pooling, social externalities, and the ability of individuals to optimize at higher levels of cost sharing, or in response to a subsidy of individuals to seek care. However, we can theoretically argue that individuals who respond to changes around higher levels of cost sharing must value the returns to care higher than at the margin of the Dutch reform. Consequently, one would expect that moral hazard is less important at those margins, and therefore the relevant importance of the downsides of cost sharing is expected to be greater than the reduction in moral hazard at these higher levels. Fifth and finally, for reasons explained previously, our strategy does not allow us to investigate the effects of cost sharing and mental health treatment for those who were not in prior treatment.

4 Analysis sample and results

4.1 Sample description

The analysis sample is constructed as follows. We take all treatment records that were opened between August 1, 2010 and July 1, 2011. We drop everyone who was younger than 17 or older than 63 years old on the date on which the record was opened that We then exclude all records that had a primary diagnosis that was removed from coverage in 2012 and 2013: adjustment disorder, other condition that may be a focus of clinical attention. In addition, we exclude dementia because this predominantly affects individuals older than 65. We also exclude records that correspond to treatment in the juvenile criminal justice system, involuntary commitment records, and records of individuals who do not appear in our baseline sample.

We investigate the smoothness of the number of daily record openings by visually inspecting Figure 5. Provider payment by insurance companies occurred after a treatment record had been closed. This meant that for records with the bulk of spending in one year, but closure in the next, costs and revenues got on the books of providers in different years. This provided an incentive to managers of provider organizations to close records at the end of a calendar year, even when treatment continued beyond that record. Given that a substantial proportion of patients were treated for multiple years, this created a cycling effect where treatment records were closed and continuation records opened in late December each year. These irregularities during the Christmas holiday also become apparent when inspecting weekly mean GAF scores, mean household income, and the proportion of these prior treatment records that was a continuted rather than an initial treatment record during November 2010 through February 2011 period. Therefore, we drop all records that were opened during the Christmas holidays in late December and early January.

[Figure 5 about here.]

Table IV shows that patients who opened a treatment record after January 1, 2011 were slightly older than those who opened a record before that date. The two groups are balanced in terms of gender, with a majority of patients being female, and in terms of GAF scores at the start of the treatment record. Those starting in 2011 were slightly more likely to have been prescribed medication, and slightly less likely to have had an inpatient stay. Average income for the 2011 subsample was .6 percent higher, while the proportion with a completed postsecondary degree was nearly the same for both subsamples.

[Table V about here.]

4.2 Results

Figure 7a shows a discontinuity around the threshold in the probability of continuing treatment for at least 30 days in the period between 13 and 18 months after opening a treatment record in the second half of 2010 or the first half of 2011. Column 1 of Panel A in Table V shows that 57.08% of individuals in our sample who opened a treatment record in 2010, and therefore didn't face the copay when continuing treatment after 365 days, did continue treatment. For individuals who opened a treatment record in 2011, conditional on the date-based running variable, this was an estimated 3.67 percentage points lower. For patients with GAF scores between one and four and relative to the baseline continuation rate, continuation rates of were an estimated 3.48% (2.46 divided by 70.64) lower for individuals who faced the copay than for those who did not. This was 6.6% for patients with a GAF score of five or six, and 4.88% for individuals with a GAF score of seven or higher. Both baseline continuation rates and the reduction in response to the copay were higher for patients who were prescribed medication; for patients in inpatient care; and for patients who were in a continued rather than an initial treatment record.

[Figure 6 about here.]

[Table VI about here.]

Figure 7b shows a reduction in the probability of being employed in any month during months 13 to 24 for patients who opened a treatment record in 2011 compared with those who opened a record in 2010. The first row of Panel A in Table VI shows that individuals who started their prior treatment record in 2010 worked an estimated 48.22% of the twelve months following the maximum duration of the treatment record. Column 4 of the same row shows that number of months in employment was reduced by an estimated 1.02 percentage points among those who faced the copay. Figure **??** shows the effect by month, estimating the model in column 4 for employment in each month separately. The point estimates decrease over time and although the confidence intervals overlap, this is consistent with the hypothesis that the absence of treatment led to a gradual depreciation of human capital that translated in reductions in employment.

[Figure 7 about here.]

To investigate whether contemporaneous shocks to employment could explain our findings. Using the matched sample, we do find a minor increase in employment which can explain at most 15.4% of the reduced-form effect.

Exploring the pathway of the reduced-form effect, we scale it by the first-stage estimate to obtain a Wald estimate of the LATE of treatment continuation on employment. We find that on average, the point estimate of the effect of treatment continuation at this margin was a 27.70% increase in the months worked out of the 12 months of the follow-up period (95% CI 4.85%–50.55%). The point estimate is largest for patients in the middle of the GAF distribution, consistent with the hypothesis that their illness was not so severe that the distance to the labor market had become too large, but that at the same time their social functioning was at a level where treatment could improve economic productivity.

[Table VII about here.]

[Table VIII about here.]

5 Conclusion

We estimate a Local Average Treatment Effect of SMHC treatment continuation on employment. The compliers are individuals who forego treatment continuation under the fixed access fee which increased out-of-pocket costs for specialist mental health care in the Netherlands. The estimates suggest that treatment continuation at this margin increases months in employment by 30%. The more seriously ill have slightly stronger estimated employment effects. The estimates for those who were unemployed in the past year are smaller for those who were employed, but still significant.

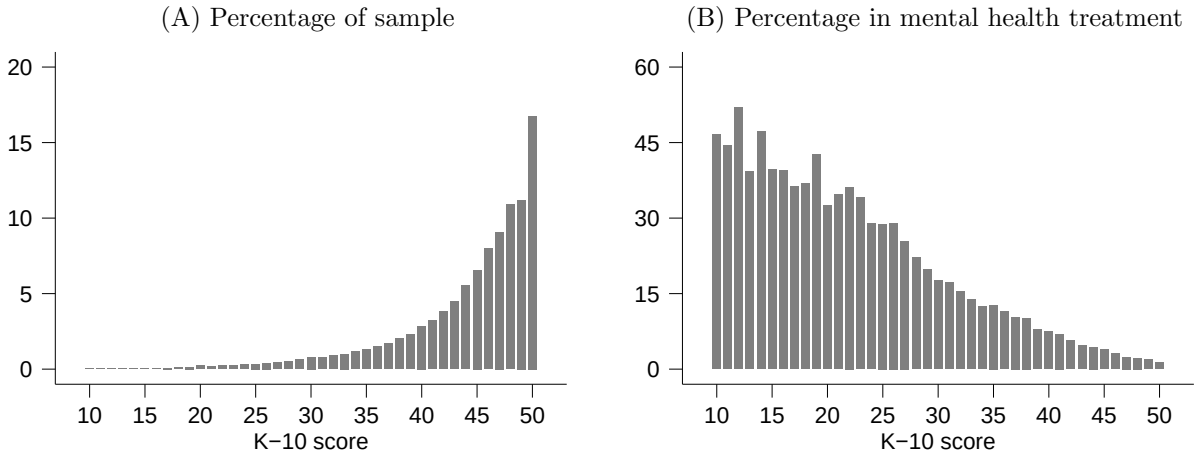

Figure 1: Distribution of K-10 scores in 2012 and proportion in treatment in 2011
Notes: Panel A shows the distribution of the K-10 scores among adults between ages 18 to 64, and Panel B shows for each separate K-10 score the percentage with that score who received mental health treatment in 2011.
Sources. GBAADRESOBJECT, GBAPERSOON, GEMON and DIS.

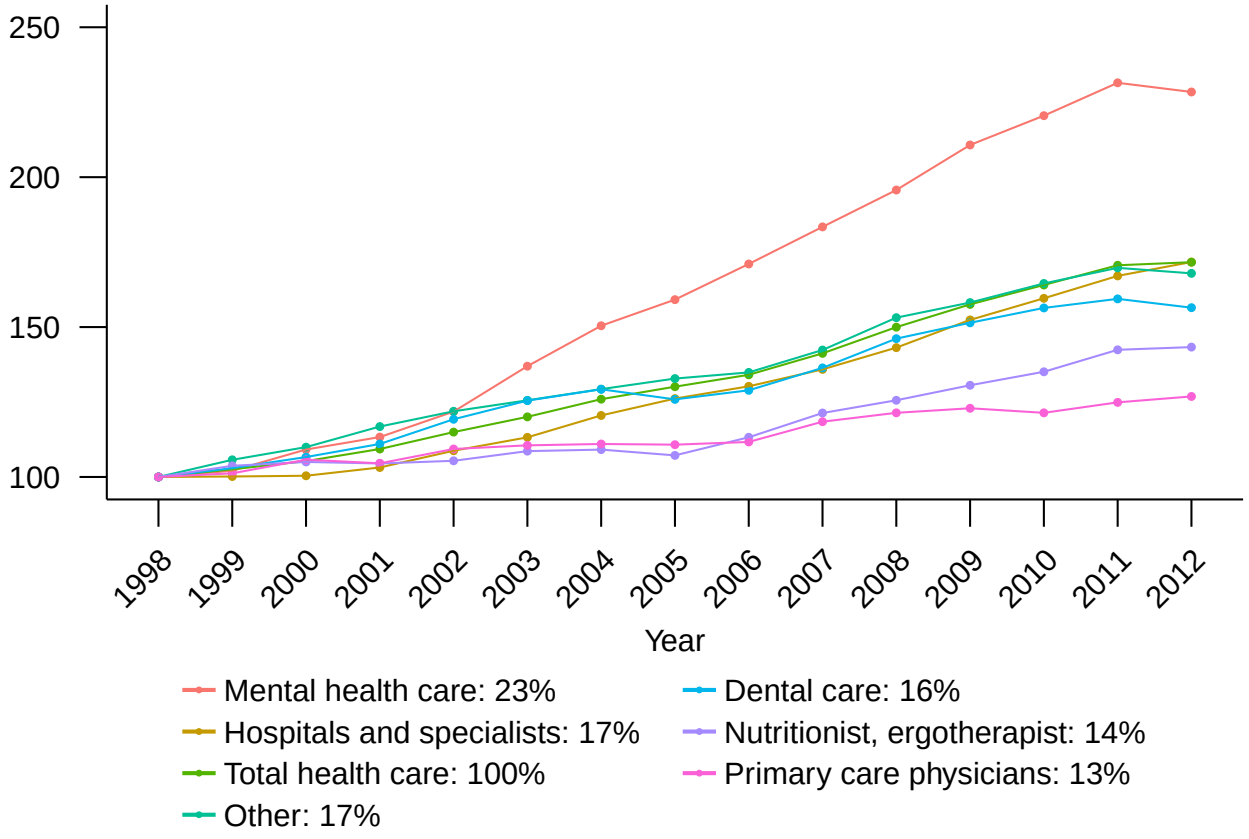

Figure 2: Normalized total annual costs by health care sector

Notes: Constructed by the authors using Dutch health account data (Centraal Bureau voor de Statistiek, n.d.).

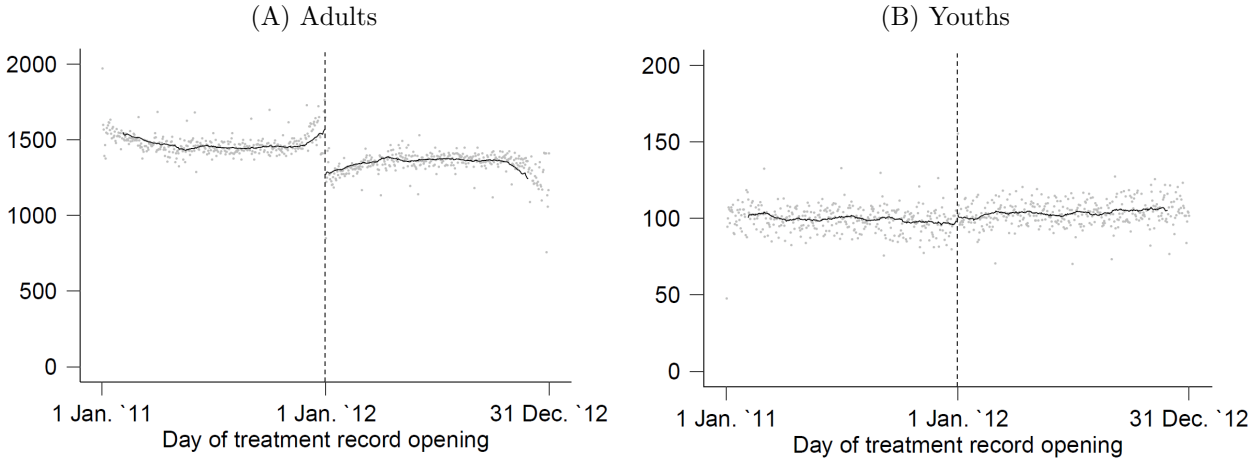

Figure 3: Daily treatment record openings for adults (left) and unaffected youths aged 15-17 (right)

Notes: Daily openings of records represented by the mean-shifted residuals of daily openings of records on day-of-the-week, day-of-the-year, and holiday indicator variables. Fitted line is constructed as the moving mean of the previous 35 days for 2010 through 2011 and as the moving mean of the subsequent 35 days for 2012. A similar figure was shown in Ravesteijn et al. (2017), but the figure presented here is based on all treatment records opened between 2010 and 2012 that could be linked to the base sample, rather than the records of only the GGZ-NL member provider organizations.

Data sources: GBA, DIS

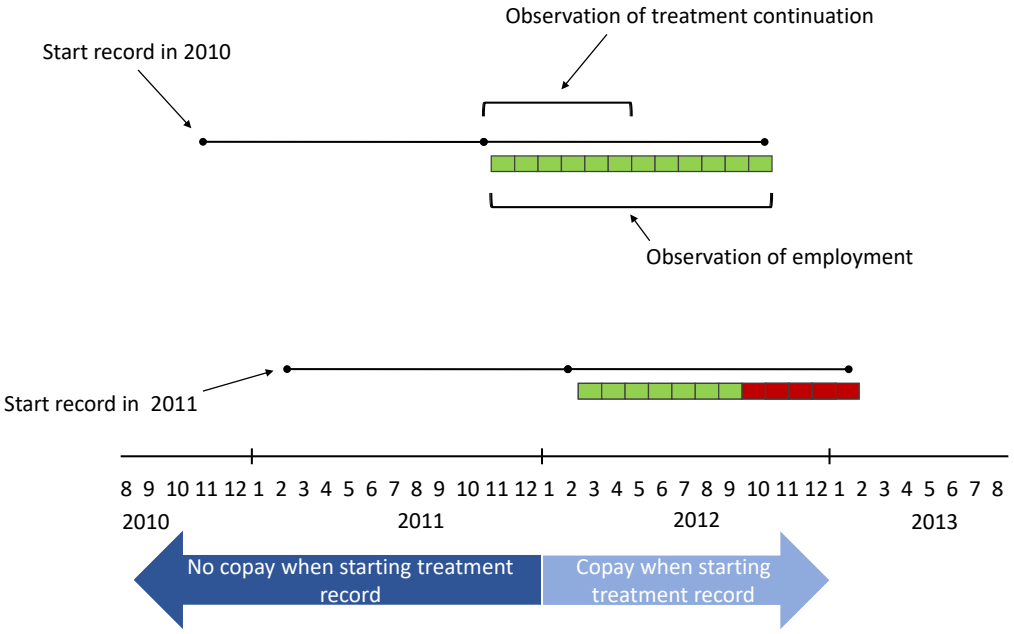

Figure 4: Schematic representation of the identification strategy

Notes: Consider two patients who open a treatment record in either November 2010 or February 2011. 365 days after opening that prior treatment record, it is automatically closed. The individual who opened the prior treatment record in 2010 can open a new treatment record without paying the copay, while the individual who opened the prior treatment record in 2011 now has to pay the copay to continue treatment. We observe treatment status in the six months after the maximum prior treatment record duration of 365 days. We refer to this follow-up period as months 13–18. An individual is considered to continue treatment if he was in treatment for at least 30 days during the six-month follow-up window. We observe employment in months 13–24, and define the proportion of months employed as the number of months with income from employment divided by twelve.

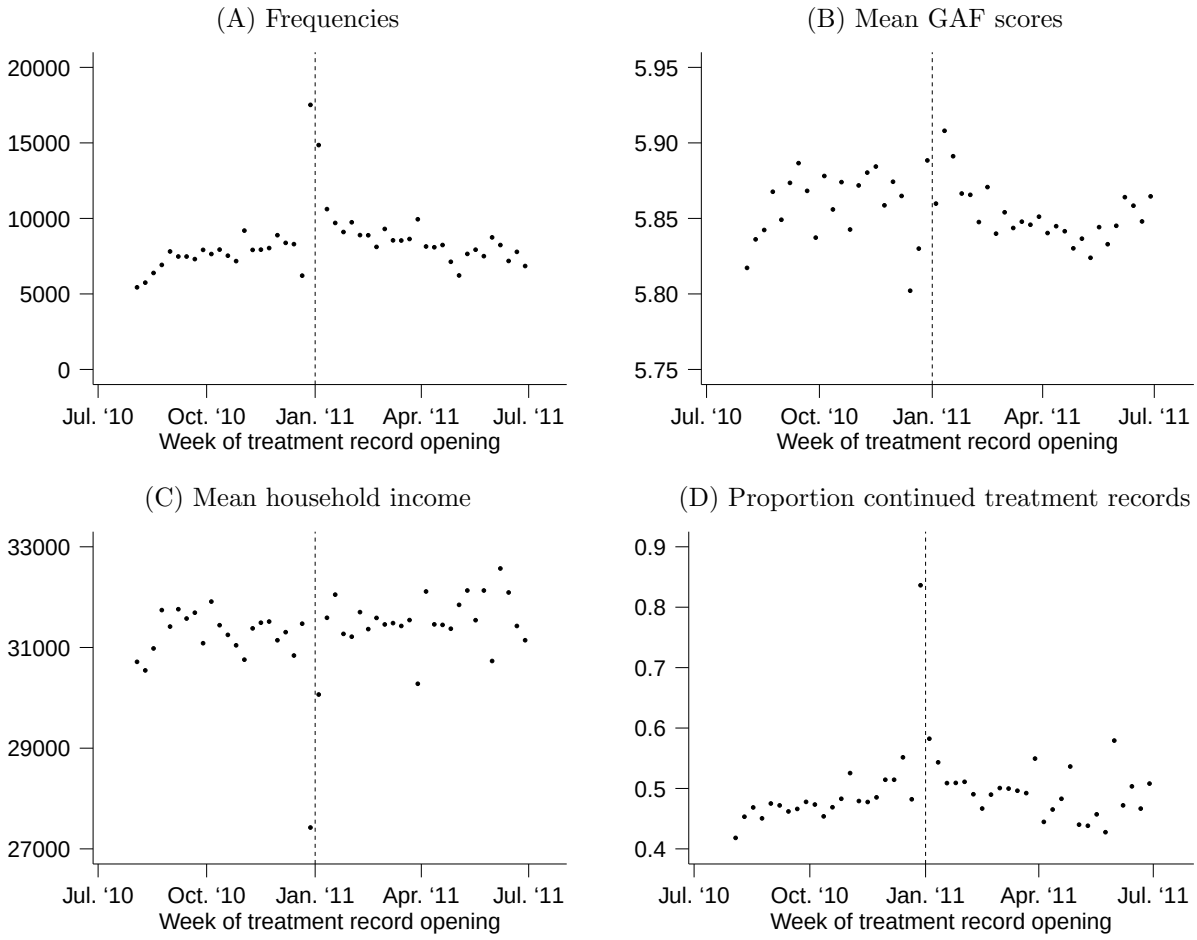

Figure 5: Patient characteristics by treatment record start date in the period around the threshold

Notes: Treatment record characteristics by start week from July 1, 2010 to June 30, 2011.
Sources: GBAADRESOBJECT, GBAPERSOON, DIS, IHI.

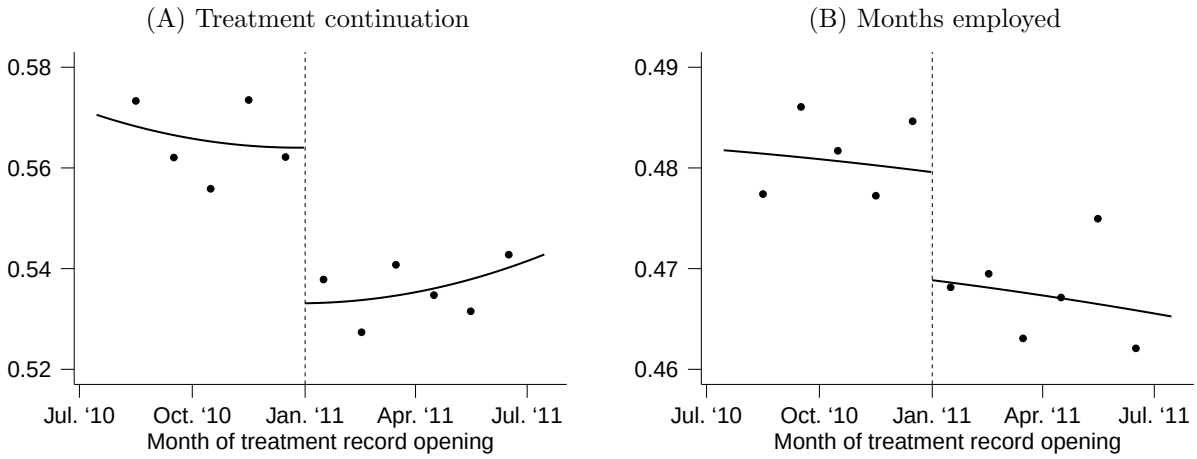

Figure 6: First stage and reduced form

Notes: Each dot represents the proportion of individuals in each month who were treatment during follow-up months 13–18 (Panel A) and the average proportion of months worked out of months 13–24 (Panel B) between August 2010 and June 2011.

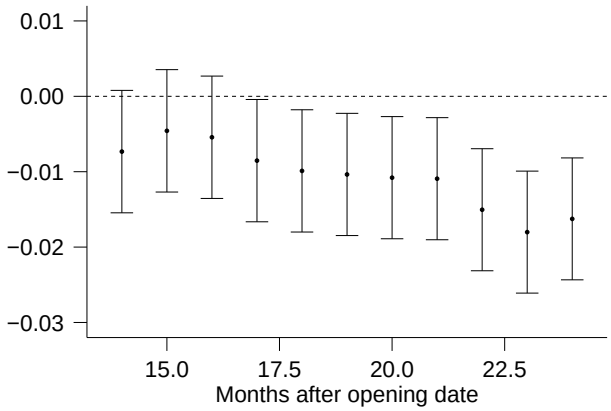

Figure 7: Estimated change in the probability of being employed (reduced-form estimate) by month

Notes: Plotted coefficients for separate reduced-form regression models (equation RF by month. Each confidence interval shows the estimated difference in the probability of being employed in each respective month by prior treatment record start year, conditional on year-of-birth dummies, the number of days since January 1, 2010, and that number squared.

Table I: Treatment record characteristics for initial and continued regular treatment, and involuntary commitment

	Initial regular	Continued regular	Involuntary commitment
Number of treatment records	634,563	517,366	1,531
Panel A: Characteristics (mean/median)			
Begin GAF	5.83	5.82	3.97
	(1.12)	(1.34)	(1.49)
Change in GAF	0.45	0.17	0.38
	(1.16)	(0.96)	(1.39)
Duration of DBC	242.28	301.85	263.36
	(130.57)	(106.64)	(141.08)
Total minutes	843	826	3,863
	[380–1,615]	[295–1,941]	[1,407–7,414]
Total visits	22.29	30.52	90.68
	(28.26)	(38.93)	(74.59)
Any inpatient days	0.08	0.09	0.82
	(0.27)	(0.29)	(0.39)
Total inpatient days if at least one	28	33	70
	[9–68]	[11–79]	[19–192]
Prescribed medication	0.27	0.50	0.65
	(0.44)	(0.50)	(0.48)
Panel B: Diagnosis (percentage)			
Adjustment disorder	9.19	6.03	0.78
Anxiety disorder	13.37	12.25	0.78
Bipolar disorder	1.91	6.24	11.37
Dementia	0.8	0.63	2.42
Depressive disorder	18.09	18.94	3.46
Disorder first diagn. in childhood	6.79	7.78	2.16
Miscellaneous disorder	8.36	5.22	1.76
Missing	10.37	0.42	0.72
Other conditions	10.68	5.72	NA
Personality disorder	8.99	12.58	4.44
Psychotic disorder	3.39	14.11	65.03
Substance-related disorder	8.06	10.07	6.8

Notes: This table shows, for initial regular treatment, continued regular treatment, and involuntary commitment records separately: the total number of treatment records for adults between 18 and 64 that were open in 2011; mean GAF scores; mean changes in GAF between the start and end of each record; the mean number of days records are open; the median total number of treatment minutes consisting of direct treatment time [interquartile range], indirect provider time, and provider travel time; the mean number of days on which treatment activities were registered; the proportion of records with any inpatient days; the median number of inpatient days for records with at least one inpatient day [interquartile range]; the percentage with recorded activity codes for prescription of pharmaceuticals. In addition Panel B shows the percentage by primary diagnosis class and reason for ending the DBC.
Data sources. GBAADRESOBJECT, GBAPERSOON, DIS.

Table II: Mean income, wealth, and hourly wage by mental health (treatment) status

	K-10 score 2012		Mental health care 2011	
	K-10≥43 mean	K-10<43 coeff	No MHC mean	MHC patients coeff
(1) Adult personal income 2011 (€)	34,224	-5,858***	30,960	-5,792***
	(31,897)	(128)	(30,151)	(34)
(2) Household income 2011 (€)	44,373	-7,232***	41,018	-9,109***
	(26,633)	(121)	(26,685)	(33)
(3) Household wealth 2011 (€)	228,391	-65,476***	171,262	-60,551***
	(830,115)	(3,766)	(908,130)	(1,173)
(4) Hourly wage (in euros)	19.52	-1.93***	18.12	-1.13***
	(16.39)	(0.09)	(17.96)	(0.03)

Notes: Columns 1 and 3 show means for those in good mental health (column 1) or those not in treatment (column 3). Columns 2 and 4 show regression coefficients for binary independent variables indicating bad mental health (column 2) or mental health treatment (column 4) conditional on year-of-birth and gender dummies, standard errors in parentheses.
Data sources: GBAADRESOBJECT, GBAPERSOON, GEMON, DIS, IPI, IHI, IVM, SPOLIS.

Table III: Employment characteristics by mental health (treatment) status

	K-10 score 2012		Mental health care 2011	
	K-10≥43 mean	K-10<43 coeff	No MHC mean	MHC patients coeff
(1) Employed 2012	0.719	−0.128***	0.724	−0.261***
		(0.002)		(0.000)
(2) Employed 2015, cond on employed 2012	0.901	−0.054***	0.898	−0.087***
		(0.002)		(0.001)
(3) Employed 2015, cond on not employed 2012	0.147	−0.047***	0.197	−0.102***
		(0.002)		(0.001)
(4) Part-time employment	0.325	0.020***	0.317	0.024***
		(0.003)		(0.001)
(5) Temporary employment contract	0.193	0.043***	0.245	0.048***
		(0.002)		(0.001)
(6) Sickness absenteeism	0.015	0.016***	0.018	0.038***
		(0.001)		(0.000)

Notes: Columns 1 and 3 show means for those in good mental health (column 1) or those not in treatment (column 3), standard deviations in parentheses. Columns 2 and 4 show regression coefficients for binary independent variables indicating bad mental health (column 2) or mental health treatment (column 4) conditional on year-of-birth and gender dummies, standard errors in parentheses. Columns 1 and 2 refer to adults aged 18-64 with a nonmissing K-10 score, columns 3 and 4 refer to all adults aged 18-64.
Data sources: GBAADRESOBJECT, GBAPERSOON, GEMON, DIS, IPI, SPOLIS.

Table IV: Mental health (treatment) status by primary income source

	% of pop (1)	K-10 score (2)	Proportion in MH treatment (3)
Employed (mean)	73.64	**44.93**	**0.045**
		(5.57)	
Unemployment insurance (coeff)	1.55	-2.11^{***}	0.062^{***}
		(0.11)	(0.001)
Disability insurance mh (coeff)	3.62	-8.92^{***}	0.314^{***}
		(0.08)	(0.000)
Disability insurance other (coeff)	5.41	-4.95^{***}	0.097^{***}
		(0.06)	(0.000)
Pension (coeff)	5.07	-0.99^{***}	0.039^{***}
		(0.06)	(0.000)
Student grant (coeff)	0.68	-1.76^{***}	0.037^{***}
		(0.23)	(0.001)
Basic welfare (coeff)	3.40	-8.61^{***}	0.198^{***}
		(0.10)	(0.000)
Other social insurance (coeff)	0.16	-3.40^{***}	0.047^{***}
		(0.31)	(0.002)
No income (coeff)	6.35	-0.93^{***}	0.006^{***}
		(0.06)	(0.000)
Overig niet actief (coeff)	0.11	-3.29^{***}	0.065^{***}
		(0.48)	(0.002)

Notes: Column 1 shows, for each panel, the percentage of individuals in each respective category. Adults aged 18–64. Income treatment status are measured in 2011, K-10 scores come from the 2012 GEMON survey. Column 1 shows the proportion of the population in each income category. The first row of each panel shows the mean K-10 score (column 2) or the proportion in treatment (3) for the baseline category (standard deviation in parentheses). All other rows show regression coefficients which indicate the differences between each respective category and the baseline category (standard errors in parentheses).
Sources: GBAADRESOBJECT, GBAPERSOON, GEMON, DIS, IPI, AOTOTUITKERING.

Table V: Characteristics of the analysis sample by prior record start year

	Start 2010 (1)	Start 2011 (2)	p-value (3)
Age	38.37	38.71	0.000
	(12.30)	(12.39)	
Male	0.45	0.45	0.319
	(0.50)	(0.50)	
Begin GAF	5.86	5.85	0.017
	(1.06)	(1.05)	
Prescribed medication	0.39	0.41	0.000
	(0.49)	(0.49)	
Any inpatient days	0.10	0.09	0.000
	(0.30)	(0.29)	
Household income 2010 (€)	27,156	27,403	0.005
	[16,088–41,224]	[16,168–41,530]	

Notes: This table shows characteristics of specialist mental health treatment records for records opened between August 1, 2010 and December 31, 2010 (column 1) and between January 1, 2011 and June 30, 2011 (column 2). Column 3 shows p-values of the coefficient of a single independent variable indicating the year of treatment initiation for each respective outcome variable in a separate OLS regression model, corresponding to each row in the table.
Sources: GBAADRESOBJECT, GBAPERSOON, DIS, IHI.

Table VI: The proportion of months 13-24 in employment for the full sample, the matched sample, and heterogeneity by characteristics of the prior treatment record

	Control mean continuing (1)	First stage (2)	Control mean employed (3)	Reduced form (4)	Wald (5)	Observations (6)
Panel A: Sample						
Full sample	0.5656	−0.0309***	0.4822	−0.0107***	0.3471**	343,406
		(0.0042)		(0.0039)	(0.1434)	
Matched sample			0.6658***	−.0000		3,803,734
			(0.0003)	(.0009)		
Panel B: Functioning score						
GAF 1–4	0.7018	−0.0178	0.2338	0.0042	−0.2383	26,603
		(0.0139)		(0.0118)	(0.6646)	
GAF 5–6	0.5823	−0.0331***	0.4529	−0.0126***	0.3807**	236,497
		(0.0050)		(0.0047)	(0.1627)	
GAF 7–10	0.4728	−0.0319***	0.6483	−0.0037	0.1165	79,874
		(0.0085)		(0.0075)	(0.2432)	
Panel C: Medication						
Not prescribed medication	0.4629	−0.0258***	0.5386	−0.0048	0.1845	205,037
		(0.0054)		(0.0050)	(0.2045)	
Prescribed medication	0.7289	−0.0549***	0.3925	−0.0104*	0.1894	138,369
		(0.0060)		(0.0060)	(0.1154)	
Panel D: Out-/ inpatient						
Outpatient	0.5432	−0.0299***	0.5063	−0.0120***	0.4024**	310,686
		(0.0044)		(0.0041)	(0.1596)	
Inpatient	0.7727	−0.0379***	0.2595	0.0016	−0.0426	32,720
		(0.0116)		(0.0110)	(0.2884)	
Panel E: Initial versus continued						
Initial	0.4964	−0.0223***	0.5499	−0.0155***	0.6955**	177,518
		(0.0058)		(0.0054)	(0.3294)	
Continued	0.6431	−0.0407***	0.4064	−0.0044	0.1072	165,888
		(0.0058)		(0.0055)	(0.1392)	

Notes: Column 1 shows, for treatment records that were opened between August 1, 2010 and December 31, 2010, the proportion of patients who continued treatment for at least 30 days in months 13–18, counting from the month in which the prior treatment record was opened. Column 2 shows regression coefficients from an OLS regression of a binary variable indicating treatment continuation for at least 30 days in months 13–18 on a binary variable indicating whether the prior treatment records was opened in 2011, conditional on year-of-birth dummies, the number of days since January 1, 2010 and that number squared. Column 3 shows, for treatment records that were opened between August 1, 2010 and December 31, 2010, the number of months divided by twelve that the individual received income from employment in months 13–24, counting from the month in which the prior treatment record was opened. Column 4 shows regression coefficients from an OLS regression the number of months divided by twelve that the individual received income from employment in months 13–24 on a binary variable indicating whether the prior treatment records was opened in 2011, conditional on year-of-birth dummies, the number of days since January 1, 2010 and that number squared. Column 5 refers to the ratio of columns 4 and 2. OLS or 2SLS standard errors in parentheses. *** $P < .01$, ** $P < .05$, * $P < .1$.

Table VII: The proportion of months 13-24 in employment, heterogeneity by employment contract characteristics in January 2011

	Control mean continuing (1)	First stage (2)	Control mean employed (3)	Reduced form (4)	Wald (4)	Observations (6)
Panel A: Employed						
Formerly employed	0.4808	−0.0284***	0.8460	−0.0090**	0.3158*	159,259
		(0.0062)		(0.0042)	(0.1701)	
Formerly not employed	0.6374	−0.0347***	0.1744	−0.0063	0.1808	184,147
		(0.0055)		(0.0039)	(0.1206)	
Panel B: Contract type						
Permanent	0.4877	−0.0229***	0.8812	−0.0053	0.2327	107,056
		(0.0075)		(0.0047)	(0.2247)	
Temporary	0.4673	−0.0408***	0.7669	−0.0181**	0.4439*	50,191
		(0.0110)		(0.0084)	(0.2540)	
Panel C: Hourly wage						
High wage	0.4653	−0.0141	0.9062	−0.0009	0.0657	81,889
		(0.0086)		(0.0048)	(0.3452)	
Low wage	0.4972	−0.0420***	0.7825	−0.0192***	0.4562**	77,370
		(0.0088)		(0.0067)	(0.1978)	
Panel D: Full time or part time						
Full time	0.4675	−0.0197**	0.8754	−0.0029	0.1463	97,273
		(0.0079)		(0.0050)	(0.2675)	
Part time	0.5016	−0.0411***	0.7999	−0.0196***	0.4768**	62,000
		(0.0099)		(0.0072)	(0.2217)	

Notes: Column 1 shows, for treatment records that were opened between August 1, 2010 and December 31, 2010, the proportion of patients who continued treatment for at least 30 days in months 13–18, counting from the month in which the prior treatment record was opened. Column 2 shows regression coefficients from an OLS regression of a binary variable indicating treatment continuation for at least 30 days in months 13–18 on a binary variable indicating whether the prior treatment records was opened in 2011, conditional on year-of-birth dummies, the number of days since January 1, 2010 and that number squared. Column 3 shows, for treatment records that were opened between August 1, 2010 and December 31, 2010, the number of months divided by twelve that the individual received income from employment in months 13–24, counting from the month in which the prior treatment record was opened. Column 4 shows regression coefficients from an OLS regression the number of months divided by twelve that the individual received income from employment in months 13–24 on a binary variable indicating whether the prior treatment records was opened in 2011, conditional on year-of-birth dummies, the number of days since January 1, 2010 and that number squared. Column 5 displays the ratio of columns 4 and 2. OLS or 2SLS standard errors in parentheses. *** $P < .01$, ** $P < .05$, * $P < .1$.

Table VIII: The proportion of months 13-24 in employment, heterogeneity by age, gender, and family migration history

	Control mean continuing (1)	First stage (2)	Control mean employed (3)	Reduced form (4)	Wald (4)	Observations (6)
Panel A: Age						
17–39	0.5430	−0.0304***	0.5326	−0.0166***	0.5467**	175,664
		(0.0058)		(0.0054)	(0.2246)	
40–64	0.5902	−0.0316***	0.4275	−0.0043	0.1374	167,742
		(0.0059)		(0.0057)	(0.1862)	
Panel B: Gender						
Female	0.5691	−0.0417***	0.4708	−0.0122**	0.2914**	188,330
		(0.0056)		(0.0052)	(0.1386)	
Male	0.5614	−0.0176***	0.4959	−0.0082	0.4626	155,076
		(0.0062)		(0.0058)	(0.3996)	
Panel C: Native						
Native	0.5605	−0.0287***	0.5167	−0.0063	0.2198	264,633
		(0.0047)		(0.0044)	(0.1660)	
Nonnative	0.5832	−0.0402***	0.3634	−0.0155**	0.3861*	78,773
		(0.0087)		(0.0078)	(0.2257)	

Notes: Column 1 shows, for treatment records that were opened between August 1, 2010 and December 31, 2010, the proportion of patients who continued treatment for at least 30 days in months 13–18, counting from the month in which the prior treatment record was opened. Column 2 shows regression coefficients from an OLS regression of a binary variable indicating treatment continuation for at least 30 days in months 13–18 on a binary variable indicating whether the prior treatment records was opened in 2011, conditional on year-of-birth dummies, the number of days since January 1, 2010 and that number squared. Column 3 shows, for treatment records that were opened between August 1, 2010 and December 31, 2010, the number of months divided by twelve that the individual received income from employment in months 13–24, counting from the month in which the prior treatment record was opened. Column 4 shows regression coefficients from an OLS regression the number of months divided by twelve that the individual received income from employment in months 13–24 on a binary variable indicating whether the prior treatment records was opened in 2011, conditional on year-of-birth dummies, the number of days since January 1, 2010 and that number squared. Column 5 displays the ratio of columns 4 and 2. OLS or 2SLS standard errors in parentheses. *** $P < .01$, ** $P < .05$, * $P < .1$.

References

Andrews, Gavin and Tim Slade, "Interpreting scores on the Kessler psychological distress scale (K10)," *Australian and New Zealand journal of public health*, 2001, *25* (6), 494–497.

Aron-Dine, Aviva, Liran Einav, Amy Finkelstein, and Mark Cullen, "Moral hazard in health insurance: do dynamic incentives matter?," *Review of Economics and Statistics*, 2015, *97* (4), 725–741.

_ , _ , and _ , "The RAND Health Insurance Experiment, Three Decades Later," 2013, *27* (1), 197–222.

Arrow, Kenneth J., "Uncertainty and the welfare economics of medical care," *Am Econ Rev*, 1963, *53*, 941–973.

Baicker, Katherine, Sendhil Mullainathan, and Joshua Schwartzstein, "Behavioral hazard in health insurance," *Q J Econ*, 2015, *130* (4), 1623–1667.

Barreca, Alan I, Melanie Guldi, Jason M Lindo, and Glen R Waddell, "Saving babies? Revisiting the effect of very low birth weight classification," *The Quarterly Journal of Economics*, 2011, *126* (4), 2117–2123.

Berwick, Donald M, Jane M Murphy, Paula A Goldman, John E Ware Jr, Arthur J Barsky, and Milton C Weinstein, "Performance of a five-item mental health screening test," *Medical care*, 1991, pp. 169–176.

Brot-Goldberg, Zarek C, Amitabh Chandra, Benjamin R Handel, and Jonathan T Kolstad, "What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics," *The Quarterly Journal of Economics*, 2017, *132* (3), 1261–1318.

Centraal Bureau voor de Statistiek, "Zorgrekeningen; uitgaven aan zorgaanbieders, financieringsbron, 1998-2013." August 15, 2017 (http://statline.cbs.nl/Statweb/selection/?VW=T&DM=SLNL&PA=81329NED&D1=a&D2=a&D3=14-15&HDR=G1%2cG2&STB=T).

Chandra, Amitabh, Jonathan Gruber, and Robin McKnight, "Patient cost-sharing and hospitalization offsets in the elderly," *Am Econ Rev*, 2010, *100* (1), 193–213.

_ , _ , **and** _ , "The impact of patient cost-sharing on low-income populations: evidence from Massachusetts," *Journal of health economics*, 2014, *33*, 57–66.

Clark, Damon and Heather Royer, "The effect of education on adult mortality and health: Evidence from Britain," *The American Economic Review*, 2013, *103* (6), 2087–2120.

Currie, Janet and Brigitte C Madrian, "Health, health insurance and the labor market," *Handbook of labor economics*, 1999, *3*, 3309–3416.

Cutler, David M and Richard J Zeckhauser, "The anatomy of health insurance," *Handbook of health economics*, 2000, *1*, 563–643.

DBC Onderhoud, "Spelregels DBC-registratie ggz RG1 Versie 20111201," Utrecht 2011.

Douven, Rudy, Minke Remmerswaal, and Ilaria Mosca, "Unintended effects of reimbursement schedules in mental health care," *Journal of health economics*, 2015, *42*, 139–150.

Dutch Healthcare Authority, "Codelijst DSM IV," Utrecht 2012.

Dutch Minister of Health, Welfare, and Sports, "Brief van de Minister van Volksgezondheid, Welzijn en Sport," July 2011.

Einerhand, Marcel and Bastian Ravesteijn, "Psychische klachten en de arbeidsmarkt," *forthcoming in ESB.*

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group, "The Oregon Health Insurance Experiment: Evidence from the First Year*," 2012, *127* (3), 1057–1106.

Frank, Richard G and Thomas G McGuire, "Economics and mental health," *Handbook of health economics*, 2000, *1*, 893–954.

Goldman, Dana and Tomas J. Philipson, "Integrated Insurance Design in the Presence of Multiple Medical Technologies," 2007, *97* (2), pp. 427–432.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," 2001, *69* (1), 201–209.

Jones, Steven H, Graham Thornicroft, Michael Coffey, and Graham Dunn, "A brief mental health outcome scale-reliability and validity of the Global Assessment of Functioning (GAF).," *The British Journal of Psychiatry*, 1995, *166* (5), 654–659.

Kelly, Mark J, Frank D Dunstan, Keith Lloyd, and David L Fone, "Evaluating cutpoints for the MHI-5 and MCS using the GHQ-12: a comparison of five different methods," *BMC psychiatry*, 2008, *8* (1), 10.

Kessler, Ronald C, Gavin Andrews, Lisa J Colpe, Eva Hiripi, Daniel K Mroczek, S-LT Normand, Ellen E Walters, and Alan M Zaslavsky, "Short screening scales to monitor population prevalences and trends in non-specific psychological distress," *Psychological medicine*, 2002, *32* (6), 959–976.

_ , Peggy R Barker, Lisa J Colpe, Joan F Epstein, Joseph C Gfroerer, Eva Hiripi, Mary J Howes, Sharon-Lise T Normand, Ronald W Manderscheid, Ellen E Walters et al., "Screening for serious mental illness in the general population," *Archives of general psychiatry*, 2003, *60* (2), 184–189.

Koning, Pierre and Maarten Lindeboom, "The rise and fall of disability insurance enrollment in the Netherlands," *The Journal of Economic Perspectives*, 2015, *29* (2), 151–172.

Manning, Willard G, Joseph P Newhouse, Naihua Duan, Emmett B Keeler, and Arleen Leibowitz, "Health insurance and the demand for medical care: evidence from a randomized experiment," *The American economic review*, 1987, pp. 251–277.

National Center for Health Statistics, "Instruction manual, part 9. ICD-10 cause-of-death lists for tabulating mortality statistics (updated October 2002 to include ICD codes for terrorism deaths for data year 2001 and WHO updates to ICD-10 for data year 2003)," Diemen 2002.

Newhouse, Joseph P. and Rand Corporation Insurance Experiment Group, *Free for all?: lessons from the RAND health insurance experiment*, Harvard University Press, 1993.

Ouwehand, P and PHM Wouters, *CAS, Classificaties voor Arbo en SV*, UWV, 2009.

Pauly, Mark V, "The economics of moral hazard: comment," *Am Econ Rev*, 1968, *58* (3), 531–537.

_ **and Fredric E Blavin**, "Moral hazard in insurance, value-based cost sharing, and the benefits of blissful ignorance," *Journal of Health Economics*, 2008, *27* (6), 1407–1417.

Ravesteijn, Bastian, Eli B Schachar, Aartjan TF Beekman, RTJM Janssen, and Patrick PT Jeurissen, "Association of cost sharing with mental health care use, involuntary commitment, and acute care," *JAMA psychiatry*, 2017, *19*.

Schneider, Silke L, "The international standard classification of education 2011," in "Class and stratification analysis," Emerald Group Publishing Limited, 2013, pp. 365–379.

Stoutjesdijk, M and E Berendsen, *De groei van de Wajonginstroom*, UWV Kenniscentrum.

Zeckhauser, Richard, "Medical insurance: a case study of the tradeoff between risk spreading and appropriate incentives," *J Econ Theory*, 1970, *2* (1), 10–26.

Appendices

A Dataset construction

In this section, we will describe how we constructed our dataset. We were able to link all data sets in this section at the individual level on the server of Statistics Netherlands.

A.1 Base sample of all Dutch residents in 2010 to 2012

The Dutch municipal basic administration person file (file name "GBAPERSOONTAB 2015") contains information on all Dutch residents and their parents between 1995 and 2015. We included only individuals who were continually registered as residing in a Dutch municipality in 2010, 2011, or 2012 using the municipal address registration file (file name "GBAADRESOBJECTBUS 2015"), which resulted in a sample of 15,791,967 individuals. Using the age on January 1, 2011, we separate the 10,172,176 adults aged 18 through 64, which we refer to as the base sample, from youths younger than 18 (3,241,990), and elderly aged 65 and older (2,377,801). Registration at the municipality of residence is mandated by "Gemeentelijke Basisadministratie Persoonsgegevens" Act. Unless indicated otherwise, all statistics in this appendix refer to the base sample.

A.2 Specialist mental health care treatment records

Since 2006, all providers were mandated by law to record information on all activities that occurred within the duration of each treatment record. These treatment records form the "DIS" dataset. Each of the treatment records contained the following variables.

Start and end dates and treatment duration Each record has information on the start and end date, which we use to construct duration of the record.

Treatment type: involuntary commitment, initial and continued regular treatment We classify each treatment record based on its care type ("Zorgtype"), a three-digit code indicating the type of treatment which is recorded within the first month of treatment(DBC Onderhoud, 2011). We identified involuntary civil commitment from these codes: "rechterlijke machtiging" (110 and 206), "rechterlijke machtiging met voorwaarden" (116 and 211), and "inbewaringstelling" (111). These three types of involuntary civil commitment are described in the Dutch law on Bijzondere Opnemingen in Psychiatrische Ziekenhuizen (BOPZ), or special commitments in psychiatric hospitals. Rechterlijke machtiging refers to involuntary commitment in a psychiatric hospital by a civil court, rechterlijke machtiging met voorwaarden refers to civil court-mandated outpatient treatment with specific conditions to prevent involuntary commitment in a psychiatric hospital, and inbewaringstelling refers to urgent involuntary commitment in a psychiatric hospital as mandated by a municipal mayor. An additional set of codes, "jeugdstrafrecht" (117 and 212), related to involuntary commitment for youths in a criminal (as opposed to civil) law framework. Because forensic mental health treatment of adults was provided in a separate provider system and not represented in our data, we dropped treatment records with either of those codes from our sample. We then classify all remaining codes between 100 and 199 as initial regular treatment, and all remaining codes between 200 and 299 as continued regular treatment.

GAF scores Global Assessment of Functioning scores describe the level of functioning on a ten-point scale and are recorded by the therapist at the beginning and end of each treatment record (Jones et al., 1995) according to the following scale:

1. "Persistent danger of severely hurting self or others (e.g., recurrent violence) OR persistent inability to maintain minimal personal hygiene OR serious suicidal act with clear expectation of death";

2. "Some danger of hurting self or others (e.g., suicidal attempts without clear expectation of death; frequent violent; manic excitement) OR occasionally fails to maintain minimal personal hygiene (e.g., smears feces) OR gross impairment in communication (e.g., largely incoherent or mute).";

3. "Behavior is considerably influenced by delusions or hallucinations OR serious impairment in communication or judgment (e.g., sometimes incoherent, acts grossly inappropriately, suicidal preoccupation) OR inability to function in almost all areas (e.g., stays in bed all day; no job, home or friends)";

4. "Some impairment in reality testing or communication (e.g., speech is at times illogical, obscure, or irrelevant) OR major impairment in several areas, such as work or school, family relations, judgment, thinking, or mood (e.g., depressed man avoids friends, neglects family, and is unable to work; child frequently beats up younger children, is defiant at home, and is failing at school)";

5. "Serious symptoms (e.g., suicidal ideation, severe obsessional rituals, frequent shoplifting) OR any serious impairment in social, occupational, or school functioning (e.g., no friends, unable to keep a job)";

6. "Moderate symptoms (e.g., flat affect and circumstantial speech, occasional panic attacks) OR moderate difficulty in social, occupational, or school functioning (e.g., few friends, conflicts with peers or co-workers)";

7. "Some mild symptoms (e.g., depressed mood and mild insomnia) OR some difficulty in social, occupational, or school functioning (e.g., occasional truancy, or theft within the household), but generally functioning pretty well, has some meaningful relationships";

8. "If symptoms are present, there are transient and expectable reactions to psycho-social stressors (e.g., difficulty concentrating after family argument); no more than slight impairment in social, occupational, or school functioning (e.g., temporarily falling behind in schoolwork)";

9. "Absent of minimal symptoms (e.g., mild anxiety before an exam), good functioning in all areas, interested and involved in a wide range of activities, socially effective, generally satisfied with life, no more than everyday problems or concerns (e.g., an occasional argument with family members)";

10. "Superior functioning in a wide range of activities, life's problems never seem to get out of hand, is sought out by others because of his or her many positive qualities; no

symptoms".

We construct a variable that measures the change in GAF by subtracting the GAF score that was recorded when a treatment record was opened from the GAF score that was recorded when a treatment record was closed.

Duration of record We create a variable measuring the total number of calendar days between the start and end dates of each treatment record.

Total minutes For each activity by a mental health professional, that professional recorded the number of direct, indirect, and travel time minutes associated with that activity. We create a variable measuring the total minutes for each treatment record by summing these three types of minutes over all activities in a treatment record.

Visits We construct a variable indicating the total number of days within the duration of a treatment record with at least one recorded activity.

Inpatient days We construct a variable counting the number of days within the duration of a treatment record for which an inpatient activity (activity codes starting with "8") was recorded.

Prescribed medication We construct a binary variable indicating whether the record contained at least one activity that involved pharmacotherapy (activity code "3.2").

Diagnosis We create twelve DSM-IV-based major primary diagnosis categories. Mental health providers in the Netherlands used a diagnostic system, known as DIS codes, to code the primary diagnosis for a treatment record. The DIS codes followed the DSM-IV structure precisely. Each code began with either "as1" or "as2" to indicate the diagnosis axis and then was followed by a sequence of up to five delimited numeric entries. Every additional numeric entry represented an added level of specificity for the diagnosis. For example, as1_6 represented mood disorder, as1_6.02 represented bipolar disorder, as1_6.02.01 represented bipolar I disorder, as1_6.02.01.03 represented bipolar I disorder where the last episode was manic, and as1_6.02.01.03.05 represented bipolar I disorder where the last episode was manic and the patient was partially in remission.

Each of the 689 observed primary diagnosis codes was an element of a main diagnostic class, which we defined as the DIS code collapsed to the first numeric entry after the axis code. For example, treatment records with primary diagnosis codes of as1_1.01.01 and

as1_1.05.01.03.01 were both elements of the as1_1 main diagnostic class. In addition, because mood disorder was such a large class consisting of two very distinct disorders, we separated it into depressive disorder (as1_6_01) and bipolar disorder (as1_6_02), and we treated each of those as its own main diagnostic class. As a result, each primary diagnosis code was an element of one of 24 main diagnostic classes.

For codes at a sufficiently high level of specificity, the Dutch Health Care Authority provided a correspondence between the DIS code and the ICD-10 classification system (Dutch Healthcare Authority, 2012). eTable1 provides for each of these main diagnostic classes all of the ICD-10 codes that correspond to any of the specific diagnoses within that class. In some cases, different DIS codes corresponded to the same ICD-10 code. We classified the primary diagnosis of the treatment record according to the DSM-IV based DIS classification as opposed to the ICD-10 classification.

DIS Code	DIS Diagnosis	ICD 10 Codes
as1_1	Disorder first diagnosed in childhood	F80.0-2, F80.9, F81.0, F81.2, F81.8-9, F82, F84.0, F84.2-3, F84.5, F84.9, F90.0, F90.9, F91.3, F91.8-9, F93.0, F94.0, F94.1-2, F94.x, F95.0-2, F95.9, F98.0-5, F98.8-9, R15
as1_2	Delirium, dementia, and amnesia	F00.xx, F01.80-81, F01.83, F01.xx, F02.00, F02.10, F02.2-4, F02.8, F03, F04, F05.5, F05.9, F06.9, F10.03, F10.4, F10.6, F10.73, F11.03, F12.03, F13.03, F13.4, F13.6, F13.73, F14.03, F15.03, F16.03, F18.03, F18.73, F19.03, F19.4, F19.6, F19.73, R41.3
as1_3	Other mental disorder due to a known physiological condition	F06.1, F07.0, F09

as1_4	Substance-related disorder	F10.03, F10.1, F10.2x, F10.4, F10.51-52, F10.6, F10.73, F10.8-9, F11.00, F11.03-04, F11.08-09, F11.1, F11.2x, F11.3, F11.51-52, F12, F12.00, F12.03-04, F12.1, F12.2x, F12.51-52, F12.8-9, F13.00, F13.03, F13.1, F13.2x, F13.3-4, F13.51-52, F13.6, F13.73, F13.8-9, F14.00, F14.03-04, F14.1, F14.2x, F14.3, F14.51-52, F14.8-9, F15.00, F15.03-04, F15.1, F15.2x, F15.3, F15.51-52, F15.8-9, F16.00, F16.03, F16.1, F16.2x, F16.51-52, F16.70, F16.8-9, F17.2x, F17.3, F17.9, F18.00, F18.03, F18.1, F18.2x, F18.51-52, F18.73, F18.8-9, F19.00, F19.03-04, F19.1, F19.2x, F19.3, F19.51-52, F19.6, F19.73, F19.8-9
as1_5	Psychotic disorder	F06.0, F06.2, F06.x, F10.51-52, F11.51-52, F12.51-52, F13.51-52, F14.51-52, F15.51-52, F16.51-52, F18.51-52, F19.51-52, F20, F20.00, F20.02-05, F20.08-09, F20.0x, F20.10, F20.12-15, F20.18-19, F20.1x, F20.20, F20.22-25, F20.28-29, F20.2x, F20.30-35, F20.38-39, F20.3x, F20.50, F20.52-55, F20.58-59, F20.5x, F20.8, F22.00, F23.80-81, F23.8x, F24, F25.0-1, F25.x, F29
as1_6.01	Depressive disorder	F32.0-4, F32.9, F32.x, F33.0-4, F33.9, F33.x, F34.1
as1_6.02	Bipolar disorder	F06.30, F06.32-33, F06.xx, F10.8, F11.8, F13.8, F14.8, F15.8, F16.8, F18.8, F19.8, F30.1-2, F30.7, F30.9, F30.x, F31.0-9, F31.x, F34.0, F39
as1_7	Anxiety disorder	F06.4, F10.8, F12.8, F13.8, F14.8, F15.8, F16.8, F18.8, F19.8, F40.00-01, F40.1-2, F41.0-1, F41.9, F42.8, F43.0-1
as1_8	Somatoform disorder	F44.4-7, F44.x, F45.0, F45.1-2, F45.4, F45.9
as1_9	Factitious disorder	F68.1
as1_10	Dissociative disorder	F44.0, F44.1, F44.81, F44.9, F48.1
as1_11	Sexual and gender identity disorder	F10.8, F11.8, F13.8, F14.8, F15.8, F19.8, F52.0, F52.10, F52.2-6, F52.9, F64.0, F64.2, F64.9, F65.0-5, F65.8-9, N48.4, N50.8, N94.1, N94.8

as1_12	Eating disorder	F50.0, F50.2, F50.9
as1_13	Sleeping disorder	F10.8, F11.8, F13.8, F14.8, F15.8, F19.8, F47.x, F51.0, F51.1, F51.2, F51.3, F51.4, F51.5, F51.8, F51.9, G47.0, G47.1, G47.3, G47.4, G47.8
as1_14	Impulse control disorder	F63.0-3, F63.8, F63.9
as1_15	Adjustment disorder	F43.20, F43.22, F43.24-25, F43.28, F43.9
as1_17	Other condition that may be a focus of clinical attention	F54, F93.3, F93.8, G21.0-1, G24.0, G25.1, G25.9, R41.8, T74.0-2, T88.7, Y04-05, Y06.1, Z55.8, Z56.7, Z60.0, Z60.3, Z61.4-6, Z61.x, Z62.4, Z63.0, Z63.4, Z63.7-9, Z71.8, Z72.8, Z76.5, Z91.1
as1_18	Additional code / no diagnosis	F99, R69, Z03.2
as1_19	Childhood disorder	NA
as2_01	Intellectual disability	F70.9, F71.9, F72.9, F73.9, F79.9
as2_16	Personality disorder	F21, F60.0-9
as2_17	Other condition that may be a focus of clinical attention	R41.8
as2_18	Additional code / no diagnosis	F99, R46.8, Z03.2
as2_19	Relationship problem	NA

After excluding these, the nine most represented main diagnostic classes opened and/or closed in 2011 (all ages): adjustment disorder; other conditions that may be a focus of clinical attention; dementia, delirium, and amnesia; depressive disorder; substance-related disorder; anxiety disorder; personality disorder; psychotic disorder; disorder first diagnosed in childhood; and bipolar disorder. To be concise, we then pooled into one "miscellaneous" category the 13 of those diagnostic classes with the fewest treatment records: other mental disorders due to a known physiological condition (4,589 records); somatoform disorder (36,402 records); factitious disorder (127 records); dissociative disorder (3,730 records); sexual and gender identity disorder (7,938 records); eating disorder (20,127 records); sleeping disorder (2,895 records); impulse control disorder (15,606 records); additional codes/no diagnosis; childhood disorder (473); and all axis 2 disorders excluding personality disorders. Finally, we created a separate category for those administrative treatment records missing a diagnosis. The resulting twelve major diagnosis categories were: depressive disorder;

substance-related disorder; anxiety disorder; psychotic disorder; personality disorder; missing; disorder first diagnosed in childhood; miscellaneous disorder; and bipolar disorder.

Closing reason Each record contained the reason for closing that record. Observed values are:

1. "Patient decision";
2. "Therapist decision";
3. "Mutual agreement";
4. "Open follow-up record";
5. "(Pre-)intake for acute care".

A.3 Mental health scores

We use three representative surveys that include measures of mental health: Gezondheidsmonitor ("GEMON"), Permanent Onderzoek Leefsituatie ("POLS"), and Gezondheidsenquête ("GECON").

GEMON is a large representative survey conducted in 2012 by Dutch municipalities among adults born before 1995, which included the 10-item Kessler Psychological Distress Scale (K-10) questionnaire (Kessler et al., 2002). 220,627 adults of working age with completed K-10 questionnaires could be linked to the GBA base sample, in addition to 124,188 elderly. This scale is used to screen for depression and anxiety, as well as for serious mental illness (Kessler et al., 2003) and contains the following items, each introduced by the question, "In the past four weeks, how often did you..."

1. "...feel tired out for no good reason";
2. "...feel nervous";
3. "...feel so nervous that nothing could calm you down";
4. "...feel hopeless";
5. "...feel restless or fidgety";
6. "...feel so restless that you could not sit still";
7. "...feel depressed";

8. "...feel that everything was an effort";

9. "...feel so sad that nothing could cheer you up";

10. "...feel worthless".

Respondents reported the duration of the response to that question, ranging from none of the time to all of the time on a five-point scale. These responses are then scored from five through to one. The minimum score is therefore 10, indicating severe distress, and the maximum score is 50, indicating no distress. There seems to be no consensus in the literature about validated cutoff points to identify mental illness (Andrews and Slade, 2001).

Second, the annual repeated cross-sectional Permanent Survey of Living Conditions (Dutch acronym POLS) was a representative survey of the full population that contained a health module ("POLSmodGEZOND") which since 2001 included the five-item Mental Health Inventory (MHI-5) questionnaire (Berwick et al., 1991). Although the POLS was discontinued after 2009, it was succeeded in 2010 by the GECON health survey, which also included the MHI-5 questionnaire and is available up to 2015.

The MHI-5 contains the following items, each introduced by the question, "How much of the time, during the last four weeks, have you..."

1. ". . . been a very nervous person?" (anxiety);

2. ". . . felt calm and peaceful?" (general positive affect);

3. ". . . felt downhearted and blue?" (depression);

4. ". . . been a happy person?" (general positive affect);

5. ". . . felt so down in the dumps that nothing could cheer you up?" (behavioral/ emotional control).

Subjects report the duration of that feeling on a six-point scale ranging from "all of the time" to "none of the time". These answers are then assigned values between 0 and 5, with "all of the time" corresponding to 5 for questions 2 and 4, and to 0 for questions 1, 3, and 5. The MHI-5 score is determined by taking the sum over the five variables, and ranges between 0 (worst mental health) and 25 (excellent mental health). In the literature, 19 has been used as a cutpoint, and scores below those thresholds define a case of common mental disorder (Kelly et al., 2008).

We observe MHI-5 in each year between 2001 and 2015, with individuals only being surveyed in one of those years such that longitudinal observation is not possible. We included

only individuals who completed all five items of the questionnaire. Annual sample sizes, conditional on a completed MHI-5 questionnaire and linkage to the GBA base file, range between 3,476 adults of working age in 2002 and 5,544 in 2014.

A.4 Monthly employment indicators

We observe, for each month between 2010 and 2015, monthly indicators of any type of employment (according to variables Xkoppelwerknsecm and/or Xkoppeldgasecm and/or Xkoppelzelfstsecm and/or xkoppelovactiefsecm) from the "SECM" file.

A.5 Employment records

From the "POLIS" (2006-2009) and "SPOLIS" (2010-2015) files we obtain highly detailed information, by month, on all employment contracts in the Netherlands. Based on these data, we construct four variables that describe employment characteristics for each employed individual in each month. Out of the 10,172,176 adults of working age in the base sample, 6,821,291 had an employment contract in January of 2011, which is the month in which we measure these variables.

Hourly wage We construct a variable measuring hourly wage by dividing the monthly gross wage for hours paid excluding overtime (variable "Sbasisuren") by the number of hours worked excluding overtime (variable "Sbasisloon").

Temporary versus permanent contract We define a temporary contract variable whether an employee is working in a temporary contract (value "B" of variable "Scontractsoort", 1,644,509 adults) versus a permanent contract (value "O" of Scontractsoort, 4,981,230 adults). We code the 195,522 observations coded as "not applicable" (value "N") as missing for this variable. A permanent employment contract in the Netherlands can typically only be terminated by the employer after payment of a relatively large severance payment or in case of bankruptcy.

Sickness absenteeism We define a sickness absenteeism indicator that measures whether an individual receives lower wages (typically a 70% cut) because employment activities were not performed due to (prolonged) illness. In January 2011, 138,270 adults (value "Z" of variable "Scdincinkverm") received lower wages because of illness, while 6,486,279 did not (value "-"). We code both the 191,803 individuals on unpaid leave (value "O") and the 4,939 individuals on pregnancy leave (value "B") as missing for the sickness absenteeism variable.

Part-time employment In the Netherlands, working 36 to 40 hours a week is considered full time, but it is common to work less than that, especially among women. We define part-time work as working 24 hours a week or less (values 1, 2, and 3 of variable "Scontractsoort"). According to this measure, 2,183,609 individuals worked part-time.

A.6 Income

We use information on both individual and household income from the Dutch Tax Office and the Dutch Social Security Administration (UWV).

Primary income source For personal income in 2011, each individual is assigned one of fourteen primary income sources (variable "Seccoal1") indicating the main source of income (see the IPI data documentation for a description of the assignment algorithm), which we collapse into ten primary income classes. First, we define the primary income class "employed" as the primary income source being employment in a private firm, employment in the government, as a CEO-owner of an organization, as being self-employed, and as being active otherwise. !he second primary income class is unemployment insurance, the third is disability insurance and Ziektewet because of mental health issues, the fourth is disability insurance and Ziektewet for other diagnoses, the fifth is pensions (regardless of age), the sixth is student grants, the seventh is basic welfare, the eight is other social insurance transfers, the ninth is no income, the tenth is "overig niet actief", and the eleventh is youths between ages 0-17.

The primary income classes are based on the fourteen primary income files, with the exception of the two DI classes. We observe all disability insurance records (data source name "AOTOTUITKERINGTAB") form the Dutch disability insurance administration (Dutch acronym "UWV") from 2007 through 2015 which contain information on claims within the four major disability insurance schemes ("WAO", "WIA", "WAZ", and both versions of the "Wajong", see Koning and Lindeboom (2015) for background information on the Dutch disability insurance system). Each record contains a variable indicating the main cause of disability "Aototdiagnosev1") according to a Dutch classification (CAS) that maps into the ICD-10 international diagnostic classification (Ouwehand and Wouters, 2009; Stoutjesdijk and Berendsen, n.d.).

We select individuals who were on DI at any point in 2011, and assign to each of them the last recorded DI diagnosis before January 1, 2012 which was reported as the cause of disability. We split these individuals into two groups: those for whom the main cause of disability was a mental or behavioral disorder (ICD-10 codes F00-F99), corresponding to CAS-codes O and P, and those for whom the cause of disability was any other diagnosis.

This "DI non mh" group includes missing diagnoses (6,224) adults after linkage to the GBA base sample) and individuals who were denoted as disability insurance and Ziektewet in variable Seccoal1, but who were not included in AOTOTUITKERINGTAB and were therefore Ziektewet recipients: 54,105 adults. Therefore, the two DI primary income classes together consist of every individual that was on DI at any point in 2011, plus individuals who were denoted as disability insurance and Ziektewet recipients according to Seccoal1 but who were not in the DI file.

Personal and household income We observe annual individual and household income from 2003 to 2015 (the respective data sources are "IPI" and "IHI". As personal income, we observe gross annual income including transfers, minus social insurance contributions but before taxes because taxes are filed jointly with the other household members. As household income, we observe annual income including transfers net of social insurance contributions and taxes (variable "Verlpersink"). Both personal and household income was top-censored at €1,000,000 and bottom-censored at €500,000. In 2011, 553 adults had top-censored and 87 had bottom-censored individual income, while 139 adults had top-censored and 254 had bottom-censored household income.

We construct a categorical variable measuring 2011 personal income. We create a truncated personal income variable for adults in 2011, separating those with primary income from student grants (770,279 individuals) and pensions (496,995 individuals) from those with strictly positive personal income from other primary income sources (8,266,152); those with negative personal income (51,271 individuals); and those with zero personal income (586,826 individuals). Income could be negative in case of a tax debt. Of adults with negative income, 89,95 percent was self-employed, recorded as one of fourteen main sources of income (see below) which was assigned to any individual who filed profit taxes. Of adults with zero personal income, 86.98 percent was female and 74.99 percent was 40 years or older, compared with only 57.55 percent of the full adult sample. We then split the truncated income variable in quartiles, with cutpoints at €18,183, €29,936, and €43,664. We will refer to these four positive income quartiles, excluding individuals who received primary income from student grants and pensions, as "positive personal income groups" because each consists of 20.32 percent of all individuals with nonmissing values in the personal income file. Our categorical personal income category has eight values: the four positive personal income groups; negative income; zero income; student grants; and pensions. The reason for separating out adults with income from student grants or pensions is that income for these individuals does not reflect lifetime income, and not taking this into account would put the majority of them in the lowest income quartile.

We then construct the categorical household variable. We separate adults with strictly positive household income (8,266,152 individuals) from those with negative (57,047 individuals) and zero (47,625 individuals) household income. Our categorical household income variable has six values: zero income, negative income, and the four positive income quartiles. All members of a household share the same household income such that the household income quartile cutpoints are based on 25th percentile of individuals, not households. The cutpoints of the positive income quartiles are €25,193, €37,228, and €51,206.

Median household income for individuals with zero personal income was €29,197 versus €38,395 for adults with strictly positive personal income. Of adults with zero personal income, only 43,057 had zero household income and 532,916 had positive household income, which indicates that most of them were living together with individuals with positive income.

A.7 Wealth

We observe detailed information on wealth from tax records in the "IVM" file for all years between 2008 and 2015. The variable "Vrlbvrw1000ver" measures the difference between assets (the sum of deposits, bonds, stocks, real estate, firm assets of self-employed individuals, and other assets in tax box 3) and liabilities (the sum of mortgage debt, other debt for consumptive purposes, financing of stocks, bonds, or annuities, financing of a second home or other real estate, and student debt) of each household. We create a categorical wealth variable indicating four positive wealth quartiles (7,786,158 adults, cutpoints at €10,802, €87,139, and €233,972), negative wealth (1,278,429 adults), and zero wealth (353,075 adults).

A.8 The municipal registries

The Dutch municipal basic administration person file (file name "GBAPERSOONTAB 2015") contains information on year and month of birth, gender, and country of birth of all Dutch residents and their parents between 1995 and 2015. Unless indicated otherwise, age is measured on January 1, 2011, and gender is measured on January 1, 2015.

Following a method used by Statistics Netherlands, we construct a variable indicating ethnicity based on the country of birth of both the parents and the individual. "Natives" are defined as individuals both of whose parents were born in the Netherlands. "First-generation migrants" are defined as individuals who were not born in the Netherlands and who have at least one parent who was born abroad, while "second-generation migrants" are defined as individuals who were born in the Netherlands and who have at least one parent who was

born abroad.

The measure of ethnicity for first- and second-generation migrants is based on the "country of origin" that was assigned to him/her according to the following rules: If at least one parent was born abroad, and the individual was born abroad as well but in a different country than any of the parents, the individual gets assigned his/her own country of birth. If the mother is born abroad, the individual is assigned as his/her country of origin the country of birth of the mother, regardless of the country of origin of the father. If only the father is born abroad while the mother was born in the Netherlands, the individual's country of origin is the country of birth of the father.

The country of origin then determines which of seven possible ethnicities is assigned. This distinction was made based on the colonial and migration history of the Netherlands: native (10,172,176 adults); other Western countries (626,308); other non-Western countries (376,234); Suriname (239,485); Turkey (229,192); Morocco (197,474); Dutch Antilles and Aruba (83,135).

Western countries were defined as all European countries excluding Turkey, North American countries excluding Mexico, all countries in Oceania, plus Japan and Indonesia. We follow the procedure of Statistics Netherlands with the exception of Indonesia, which we define as a separate eighth ethnicity because of the large proportion if Dutch residents who descend from Indonesian migrants, many of whom came to the Netherlands after the independence of Indonesia (292,308). Our reason for doing this is that we want to study the mental health of these individuals separately given the potential impact of colonization by the Dutch and the Japanese, World War II and the subsequent Indonesian War of Independence.

A.9 Educational attainment

Information on educational attainment of 5,295,690 adults was taken from the 2011 Dutch educational attainment database (file name "HOOGSTEOPLTAB2011"), which combines information on enrollment and graduation from educational registers with self-reported educational attainment in a rotating panel survey among the labor force (file name "EBB") and administrative records from government organizations that execute active labor market policies (Dutch acronym UWV). This database contains all followed and completed educational qualifications after 2006, and a nonrandom subset before that year, such that the data is only complete for younger birth cohorts who had completed their final educational qualification after 2005. Sampling for older cohorts could be related with mental health status, for example because older individuals who were not in the labor force from a young age were not sampled by the labor force survey. This should be kept in mind while interpreting

our results.

To map each of the 23,681 Dutch educational codes (variable name "Oplnr") in our data into nine ISCED 2011 levels, we use two crosswalks that were provided by Statistics Netherlands. First, we use the "OPLEIDINGSNRREV15" file to map "Oplnr" into the "Cto" variable, which left us with 103 unique educational codes. Next, we use the "CTOREFV6" file to map "Cto" into variable "Isced2011levelhb", which consists of nine ISCED 2011 educational levels ranging from pre-primary education and development to doctoral or equivalent level (Schneider, 2013). We then created a tenth category for individuals who were 17 years old or younger on January 1, 2011.

A.10 Household type

The municipal basic administration household file (file name GBAHUISHOUDENSBUS) contains information about the position of each individual in the household, with information on five percent of the population coming from tax records and basic welfare records. We construct a variable indicating an individual's position in a household on January 1, 2011, based on the ten household position values of variable "Plhh". The resulting seven categories are: a child living at home; a single adult; a member of a couple without children; a member of a couple with children; a single parent; someone living in an institutional household (e.g. nursing homes, children's homes, rehabilitation centers, or penal institutions); or other. A couple can have either a legal status (marriage or registered partnership) or refer to a cohabiting couple without a legal contract.

A.11 Mortality

From the cause-of-death registries we obtained information on the cause of death between 2010 and 2015 according to the ICD-10 classification. We constructed six binary indicators of mental health-related causes of death, and a seventh indicator for all other causes of death. Based on the observed codes in the data and a manual of the U.S. Department of Health and Human Services (National Center for Health Statistics, 2002), we defined suicide as ICD-10 codes X60-84; poisoning as X40-49 and Y10-19; homicide as X90-99, Y0-9, and Y87.1; dementia and other organic mental disorders as F0-9; mental and behavioral disorders due to psychoactive substance use as F10; other mental illnesses as F20-F99; and any other cause of death.

A.12 Health care claims

Health insurers are mandated to send information on all claims to a central data warehouse (Dutch acronym "VEKTIS"). For each health insurance enrollee and each year between 2009 and 2014, we observe total annual payments for each of sixteen types of health care that are considered essential health benefits by the ZVW (file name "ZVWZORGKOSTENTAB". We observe the 2011 claims of 10,056,158 adult enrollees. We keep separate spending on specialist mental health care, hospital care, pharmaceutical drugs, and primary care physician care. From the latter category, we subtract the element "subscription costs" which is what the large majority of Dutch resident pay to their primary care physician regardless of use. 9,937,300 adults have positive PCP subscription claims, 99 percent of PCP annual subscription claims are below €73.88. We pool all remaining spending categories: dental care; paramedic care (e.g. physical therapy insured within the basic coverage package of essential health benefits, speech therapy); medical devices (e.g. crutches, hearing devices); patient transport; natal care; health care abroad; primary mental health care; and a variable indicating health care costs covered according to the ZVW that are not contained in the other variables ("other claims"). Claims for health services that were paid for by supplemental insurance are not part of the data. We then pool all health spending minus the PCP subscription fee to create a variable that measures total claims in euros.

We observe the Individuals could have negative claims as a result of year-to-year adjustments. Very few adults had negative claims, indicating an adjustment for previous years: 1,496 individuals for the "other", 1,467 for hospital claims, and fewer for the remaining claims categories. 73,903 adults had zero total claims; 9,470,700 had zero specialist mental health care claims; 4,230,961 had zero hospital claims; 2,850,060 had zero pharmaceutical claims; 111,718 had zero PCP claims; 6,648,309 had zero other claims.

We then split each claims category at the median of positive claims, which is €404,22 for total claims; €1,920.71; 453,97 for hospital claims; €95.10 for pharmaceutical claims; €43.37 for pcp claims; and €155.00 for other claims.

Specialist mental health care claims were billed after a treatment record was closed. Therefore, claims that are billed in a given year could refer to treatment in that same year, or in the previous year. In the claims file, total specialist mental health care spending is separated into four cost types: inpatient care up to one year; outpatient care by provider organizations; outpatient care by self-employed mental health care providers; and other care paid for by the ZVW, e.g. acute care and methadone provision.

A.13 Long-term care

The data file that contains all long-term care records ("ZORGMVTAB") contains start and end dates of each institutionalization in accordance with the AWBZ act, and separate variable specifying the so-called care intensity package which contains information on in which of the three types of long-term care institution the patient was institutionalized. Values 19900-20100 correspond to nursing homes, values 30270-30790 correspond to homes for the disabled, and values 40810-41050 correspond to mental health care institutions.

Our long-term care variable indicates whether an individual was in long-term care at any point in 2011, and if so in which type of institution: nursing homes (29,517 adults, note that these are all under 65 years old), homes for the disabled (45,837 individuals), mental health care institutions after the first year of admission (21,919 individuals); and unclassified institutions (7 individuals). For individuals who were in more than one type of long-term care in 2011, we took the type with longest total stay. Ties, which we observed between nursing homes and homes for the disabled, were assigned to nursing homes. We then create a variable indicating long-term care use in 2013, conditional on not being in long-term care in 2013: 6,072 individuals in nursing homes, 4,937 in disability homes; 5,513 in mental health homes; 11 in unclassified institutions.

A.14 Data linkage

Tables A.I–A.III show the link rates for each of the variables.

[Table A.I about here.]

[Table A.II about here.]

[Table A.III about here.]

Table A.I: Linkage of treatment record information

	Observations in data	With nonmissing value	In GBA 2010–2012	Of working age
Start date	2,842,430	2,842,430	2,685,069	1,842,155
End date	2,842,430	2,842,430	2,685,069	1,842,155
Begin GAF	2,842,430	2,596,848	2,464,863	1,697,438
End GAF	2,842,430	2,580,567	2,449,378	1,686,862
Total minutes	2,842,430	2,841,957	2,684,691	1,841,867
End reason	2,842,430	2,842,430	2,685,069	1,842,155
Primary diagnosis	2,842,430	2,842,430	2,685,069	1,842,155

Table A.II: Linkage of administrative data

	Observations in data	With nonmissing value	In GBA 2010–2012	Of working age	% of base sample
Month of birth	21,678,831	21,678,831	15,791,967	9,931,758	100.00
Gender	21,678,831	21,678,831	15,791,967	9,931,758	100.00
Ethnicity	21,678,831	21,678,831	15,791,967	9,931,758	100.00
Generation	21,678,831	21,678,831	15,791,967	9,931,758	100.00
Household position 2011	16,143,433	16,143,433	15,362,711	9,685,427	97.52
Educational attainment	9,381,831	9,381,831	8,804,836	5,389,574	54.27
Contract type January 2011	7,527,467	7,527,467	7,145,589	6,848,734	68.96
Sickness absenteeism January 2011	7,527,467	7,527,467	7,145,589	6,848,734	68.96
Contract wage January 2011	7,527,467	7,527,467	7,145,589	6,848,734	68.96
Contract hours January 2011	7,527,467	7,527,467	7,145,589	6,848,734	68.96
Contract hours class	7,527,467	7,527,467	7,145,589	6,848,734	68.96
Primary income source 2010	16,691,360	16,691,360	15,790,785	9,930,817	99.99
Primary income source 2011	16,766,891	16,766,891	15,791,197	9,931,144	99.99
Primary income source 2012	16,779,575	16,779,575	15,791,258	9,931,199	99.99
Primary income source 2015	16,900,726	16,900,726	15,384,023	9,779,214	98.46
Cause of disability 2010	874,721	874,721	809,457	698,124	7.03
Cause of disability 2011	877,063	877,063	820,652	740,645	7.46
Personal income 2011	16,730,348	16,730,348	15,791,162	9,931,111	99.99
Disposable household income 2011	16,730,348	16,730,348	15,791,162	9,931,111	99.99
Household assets 2011	16,730,348	15,487,929	14,756,330	9,172,447	92.35
Cause of death 2013	141,245	141,245	139,453	23,117	0.23
Cause of death 2014	139,223	139,223	137,276	24,302	0.24
Total health care claims 2011	16,831,682	16,491,149	15,656,123	9,816,356	98.84
Hospital claims 2011	16,831,682	16,491,149	15,656,123	9,816,356	98.84
Mental health care claims 2011	16,831,682	16,491,149	15,656,123	9,816,356	98.84
Pharmaceutical claims 2011	16,831,682	16,491,149	15,656,123	9,816,356	98.84
PCP claims 2011	16,831,682	16,491,149	15,656,123	9,816,356	98.84
Other claims 2011	16,831,682	16,491,149	15,656,123	9,816,356	98.84
Long-term care 2013	112,104	112,104	109,523	15,896	0.16
In GBA 2010–2012	15,791,967	15,791,967	15,791,967	9,931,758	100.00

Table A.III: Linkage of survey data

	Observations in data	With nonmissing value	In GBA 2010–2012	Of working age	% of base sample
MHI-5 2001	9,594	5,689	4,484	3,389	0.03
MHI-5 2002	9,667	5,561	4,470	3,318	0.03
MHI-5 2003	9,804	6,262	5,090	3,825	0.04
MHI-5 2004	11,030	7,253	5,969	4,585	0.05
MHI-5 2005	10,303	6,761	5,609	4,262	0.04
MHI-5 2006	9,607	6,413	5,374	4,103	0.04
MHI-5 2007	8,740	5,636	4,811	3,570	0.04
MHI-5 2008	9,498	5,925	5,099	3,682	0.04
MHI-5 2009	9,118	5,598	4,867	3,453	0.03
MHI-5 2010	16,337	7,019	6,874	4,871	0.05
MHI-5 2011	14,453	6,115	6,029	4,239	0.04
MHI-5 2012	14,349	6,307	6,234	4,395	0.04
MHI-5 2013	13,744	6,366	6,292	4,453	0.04
MHI-5 2014	9,516	7,994	7,849	5,341	0.05
MHI-5 2015	9,358	7,927	7,783	5,330	0.05
K-10 2012	387,193	350,417	346,834	199,703	2.01

B Associations between mental health (treatment) and individual characteristics

B.1 Treatment record characteristics

[Table B.I about here.]

[Table B.II about here.]

B.2 Mental health and treatment

We investigate the relationship between mental health scores, demographic characteristics, and health care spending. Panels A, C, and D of Figure B.3 shows the distributions of Kessler-10 and Mental Health Inventory-5 scores for the both samples. High scores indicate good mental health. MHI-5 scores were pooled into five categories to keep cell sizes above ten observations, which is required by Statistics Nederlands

27.83 percent percent of MHI-5 scores of adults, pooled across the 2001-2015 waves, were below the threshold of 19. We construct a cutpoint for the K-10 score by using scores that match closest with the percentile of the MHI-5 cutpoint. 27.72 percent of adults had K-10 scores below 43, which we will use as a cutpoint to split the sample in good and bad mental health.

In our data, 984,476 individuals had an open treatment record at any time in 2011, while 14,653,701 did not. Panels B, D, and F of Figure B.3 show the proportion of individuals in treatment by mental health score. Of individuals with K-10 scores between 10 and 15 in 2012, 36.27 percent was in treatment in 2011.

K-10 scores come from a large sample (GEMON) but were measured in 2012, while in this descriptive section, we look at treatment in 2011 which was the year before the cost-sharing reform. To investigate whether mental health scores in 2012 can be used to explain need for mental health care use in 2011, we perform the same analyses for the two smaller cross-sectional samples of MHI-5 scores in 2011 and 2012. Panels D and F of Figure B.3 reveal that MHI-5 scores in both years predict mental health care use in 2011.

[Figure B.1 about here.]

Panel A of Figure B.2 shows average MHI-5 scores of individuals who had an open treatment record in 2011 versus those who were not, for each survey year between 2001 and 2015. Panel B shows the proportion with an MHI-5 score of 19 or higher. This figure shows

how, on average, there is a downward trend in MHI-5 scores up to 2011, while scores improve after 2011, albeit still being much lower than the scores of those who were not in treatment in 2011.

[Figure B.2 about here.]

[Table B.III about here.]

B.3 Mental health (care), economic productivity, and social insurance

[Figure B.3 about here.]

[Figure B.4 about here.]

[Table B.IV about here.]

Table B.V shows the association of mental health and treatment status with income. Note that Panel A of this figure coincides with Figure IV in the main text. Not surprisingly, being on social insurance is associated with worse mental health and a higher probability of receiving mental health care. What is surprising however, is that individuals who are on disability insurance because of mental illness (DIMI claimants) only have an estimated 30.81 percentage point increase in the likelihood of receiving specialist mental health care compared with employed individuals. This is remarkable because it means that more than half of DIMI claimants are not being treated for in in the specialist mental health care sector. Previous work has shown that 46.5% of DIMI claimants uses psychotropic pharmaceuticals, and that in total 52.8% of DIMI claimants use specialist mental health care and/or psychotropic pharmaceuticals in 2013 (Einerhand and Ravesteijn, n.d.). This leaves 148,625 of these DIMI claimants without either of these types of care. Data on the type of provided pharmaceuticals is not yet available to us at this time.

Out of 10,056,400 working-age individuals between 18 and 65 years old, 13% (1,301,488 individuals) received income support in the form of UI, DI, basic welfare, or other social insurance. However, these individuals were responsible for 58% of specialist mental health care costs (€1.50 billion out of a total of €2.57 billion).

While individuals who are unable to work because of mental illness are eligible for DI, mental health scores for individuals on basic welfare, which has lower payments than DI, are comparable to scores of DIMI claimants (column 3 of Panel A of Table IV). The likelihood of

receiving specialist mental health care for basic welfare recipients is 19.85 percentage points higher than for employed individuals.

Income quartiles are conditional on not having zero or negative income, or having student grants and pensions as the primary income source in 2011. If we would not separate these latter two categories out, most of these individuals would be in the lowest income quartile on the basis of their income in 2011, and K-10 scores would be lower for the second than for the first income quartile, while the percentage in MH treatment would be higher.

The pension category in column 1 of Panel B of Table IV is slightly larger than in Panel A because it contains some individuals who also had DI claims in 2011, while these were coded as one of the two DI groups in Panel A. Of all individuals in DI, 43% were awarded DI because of mental illness.

[Table B.V about here.]

Table B.VI shows the association of mental health and treatment status with household income and wealth. Panel A shows a monotonic gradient between household income and mental health scores, with a 42.01 average K-10 score for the lowest household income quartile, while the highest income quartile has a predicted score that is 3.53 higher, conditional on age and gender.

12.1% of adults in the lowest household income quartile are in mental health treatment, while this is a predicted 8.4 percentage points higher (conditional on age and gender) for those in the highest income quartile. Panel B of Table B.VI shows a similar pattern by household wealth.

[Table B.VI about here.]

B.4 Mental health (treatment) and demographic characteristics

Figure B.5 shows that mental health scores cross-sectionally improve after adolescence, stabilize or even slightly decrease among individuals between 40 and 55 years old, and then gradually increase. We are not able to determine whether this is due to improved mental health, selective mortality, or differences in reporting behavior. After the age of 67 mental health steadily decreases. MHI-scores (not shown here) display a similar pattern. Because we observe MHI-5 scores for youths, we can report that these scores are higher throughout childhood but substantially decrease right before adolescence.

[Figure B.5 about here.]

[Figure B.6 about here.]

[Table B.VII about here.]

[Table B.VIII about here.]

[Table B.IX about here.]

B.5 Mortality, health care claims and long-term care use

[Table B.X about here.]

[Table B.XI about here.]

[Table B.XII about here.]

[Table B.XIII about here.]

[Figure B.7 about here.]

[Figure B.8 about here.]

[Figure B.9 about here.]

[Figure B.10 about here.]

[Figure B.11 about here.]

[Figure B.12 about here.]

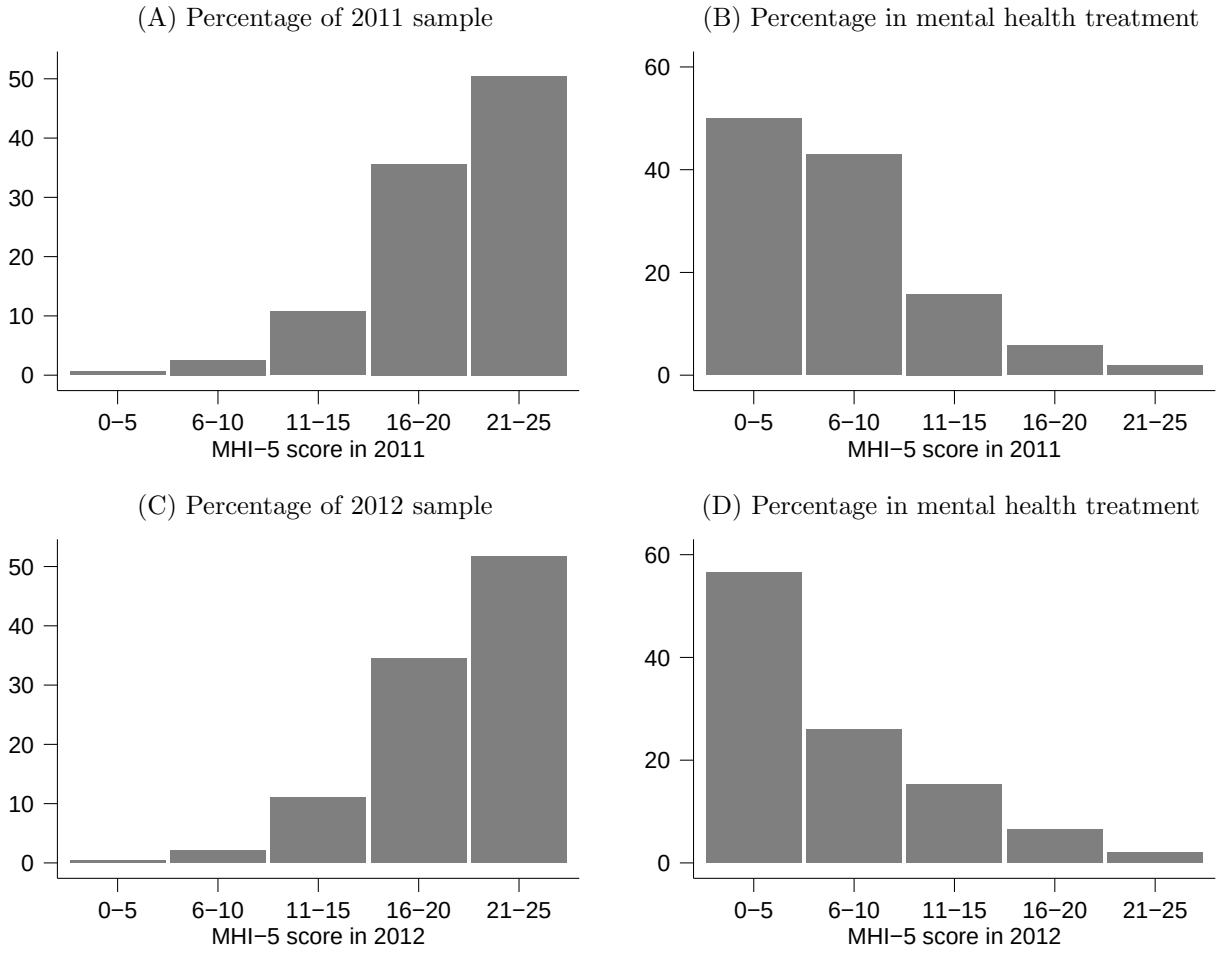

Figure B.1: Distribution of MHI-5 scores and proportion in treatment in 2011

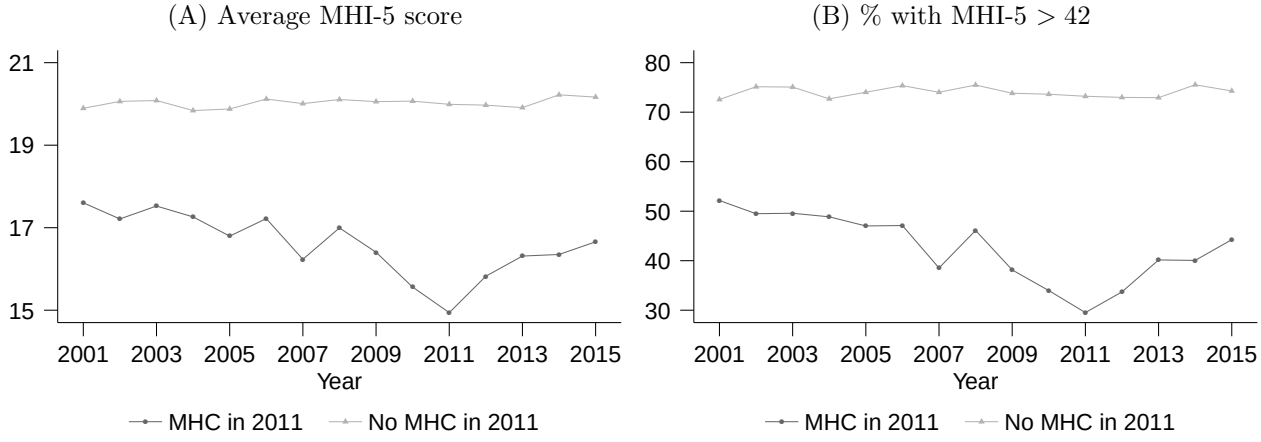

Figure B.2: MHI-5 scores before, during, and after treatment by treatment status in 2011

Notes: MHI-5 scores are measured in cross-sectional surveys (POLS and GECON).

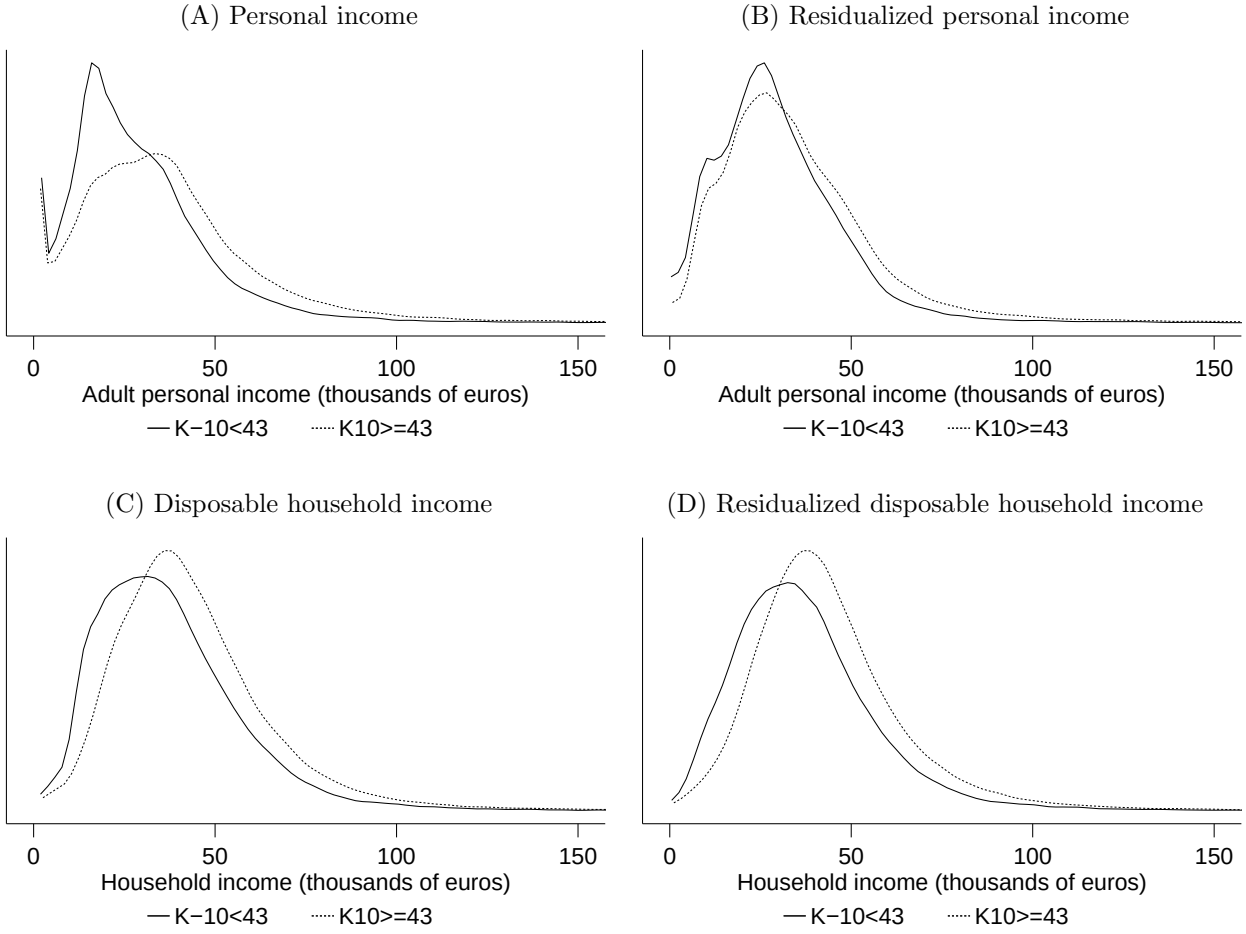

Figure B.3: Income by mental health status

Notes: Kernel smoothed densities for adults between 18 and 64 years old. Mental health scores are measured in GEMON in 2012, K-10¡43 refers to poor mental health, K10¿=43 refers to good mental health. Income is measured in 2011. Panels A and C show the smoothed density of personal and disposable household income, respectively, by mental health status. Panels B and D show the smoothed density of personal and disposable household income by mental health status, conditional on gender and year-of-birth dummies. Income is bottom-censored at zero.

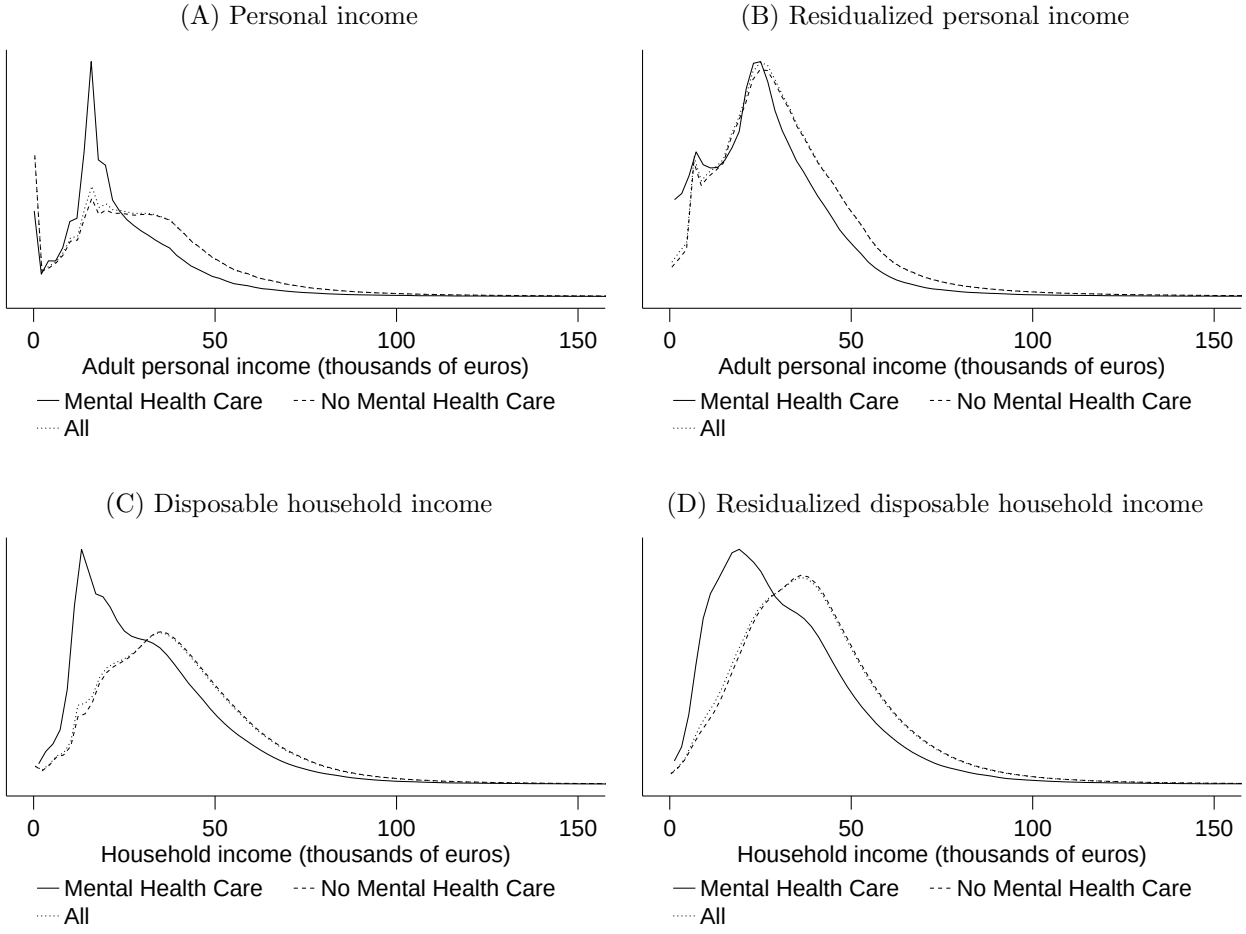

Figure B.4: Income by mental health treatment status

Notes: Kernel smoothed densities for adults between 18 and 64 years old. Panels A and C show the smoothed density of personal and disposable household income, respectively, by mental health treatment status. Panels B and D show the smoothed density of personal and disposable household income by mental health treatment status, conditional on gender and year-of-birth dummies. Income is bottom-censored at zero.

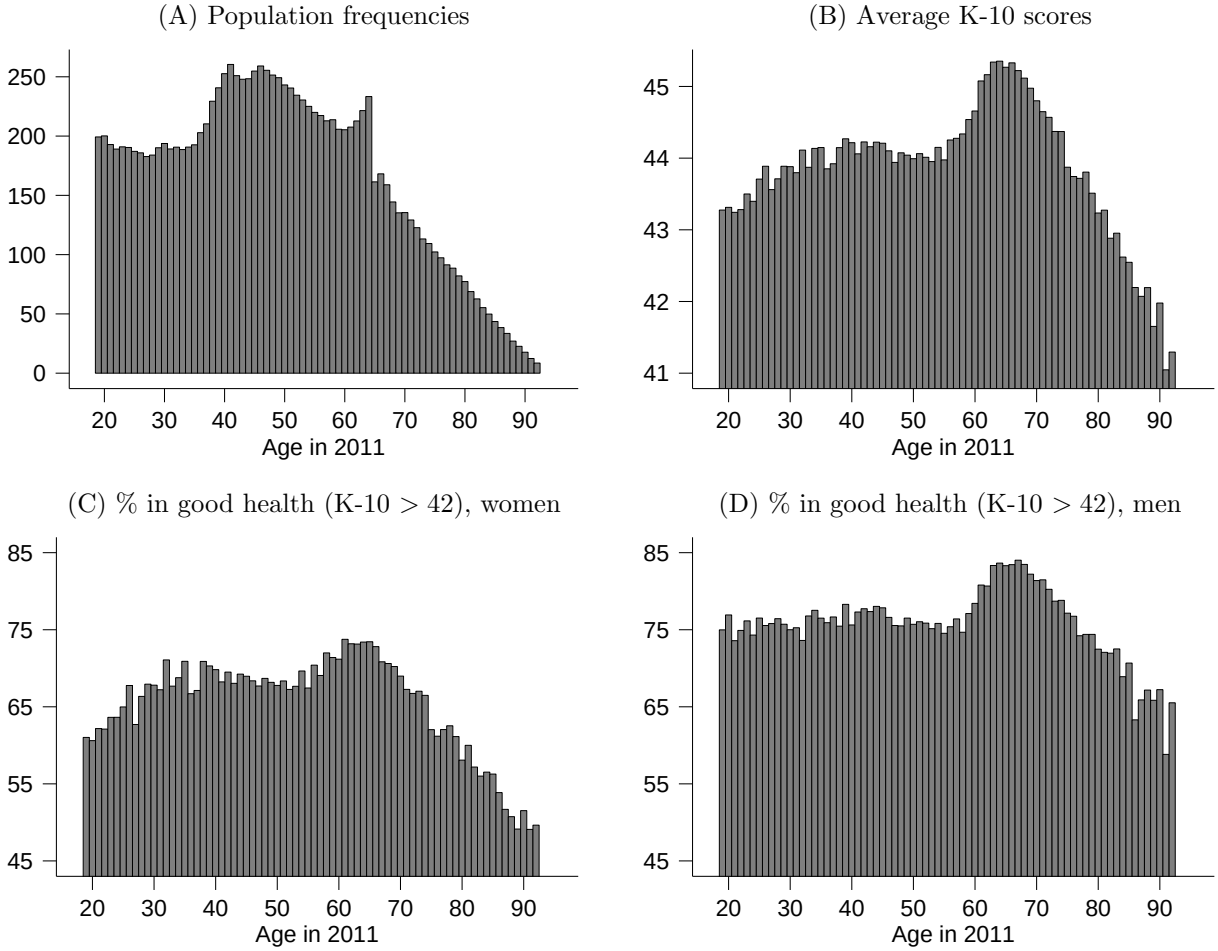

Figure B.5: Mental health scores over the life cycle

Notes: K-10 scores are measured in 2012.

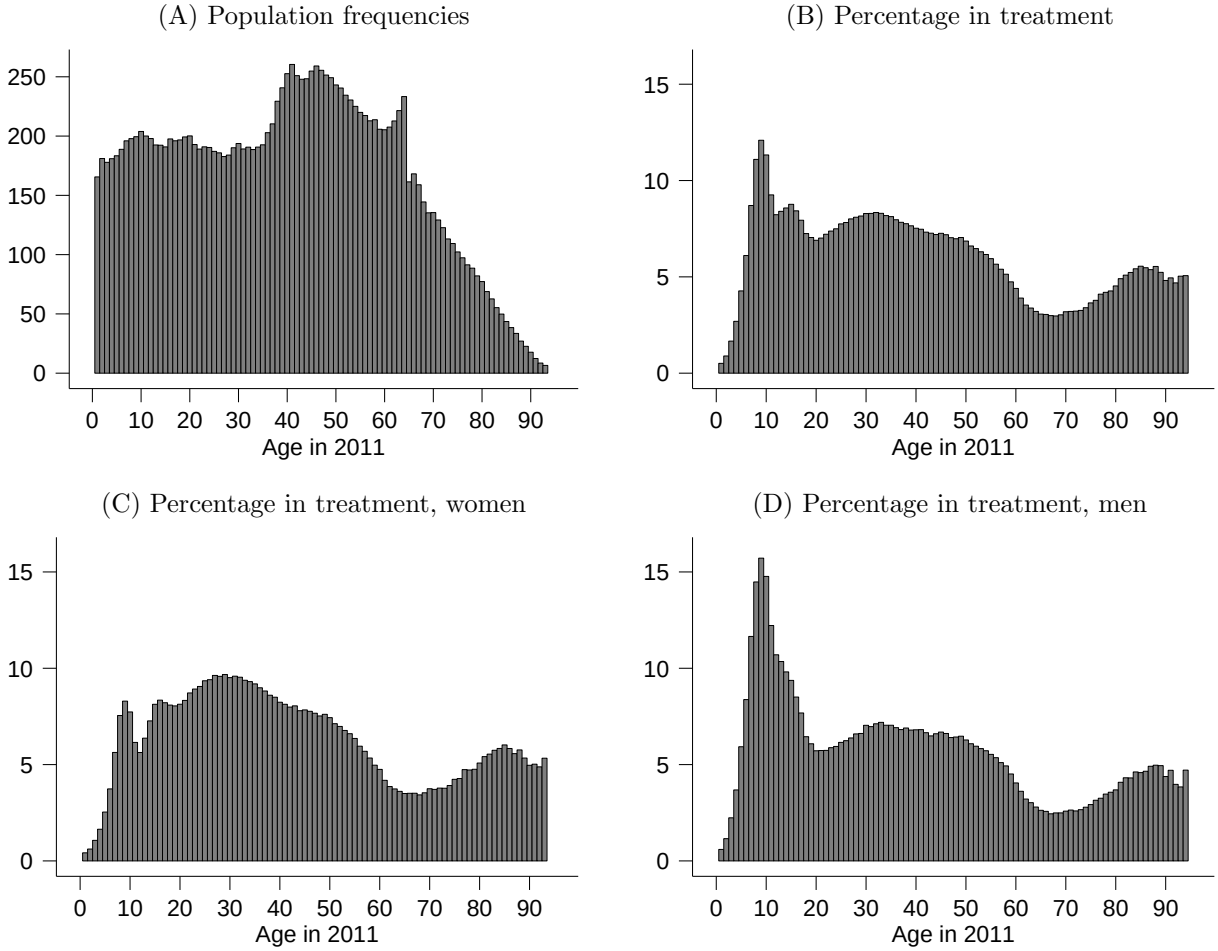

Figure B.6: Mental health treatment over the life cycle

Notes: Treatment status is measured in 2011.

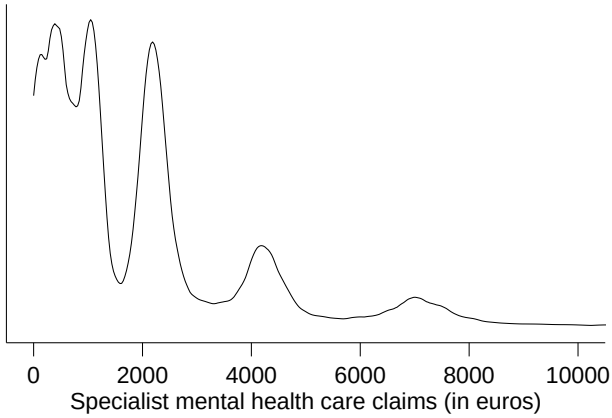

Figure B.7: Smoothed probability density function of specialist mental health claims in 2011

Notes: The bunching is caused by disconinuities in the tariff schedule (see Douven et al., 2015)

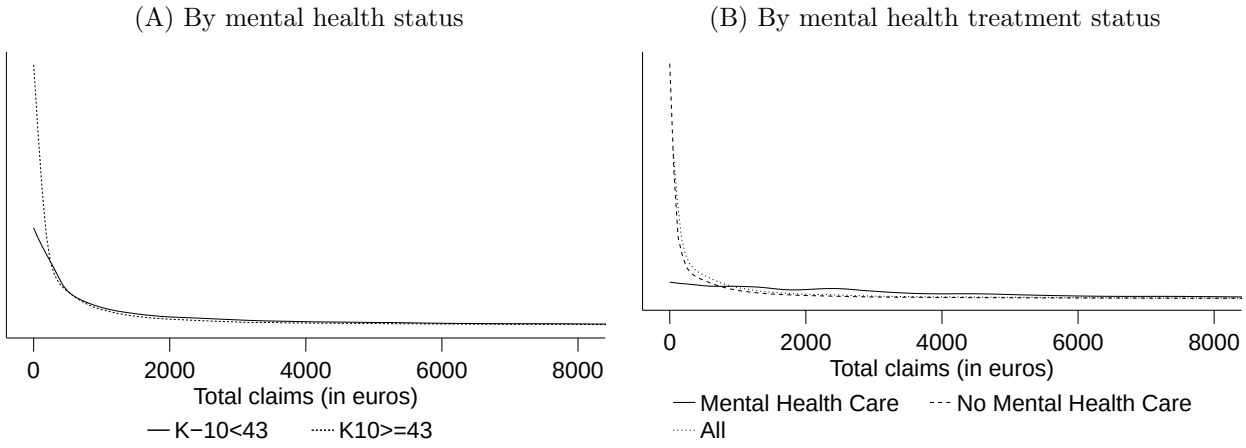

Figure B.8: Total health care claims in 2011 minus PCP subscription fees

Notes: The figures show smoothed probability density functions. K-10 scores were measured in the 2012 GEMON survey, treatment status was assessed in 2011 for the full population.

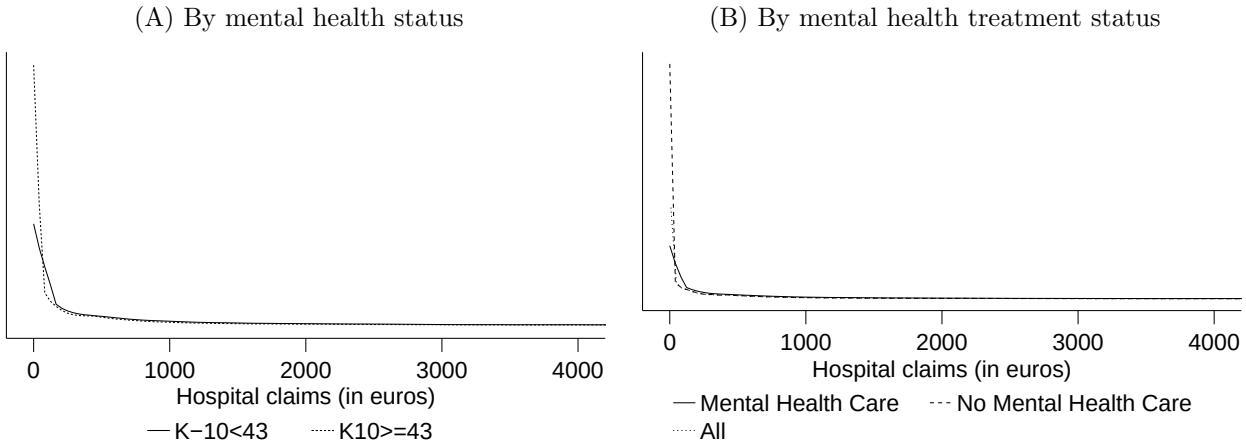

Figure B.9: Hospital claims in 2011

Notes: The figures show smoothed probability density functions. K-10 scores were measured in the 2012 GEMON survey, treatment status was assessed in 2011 for the full population.

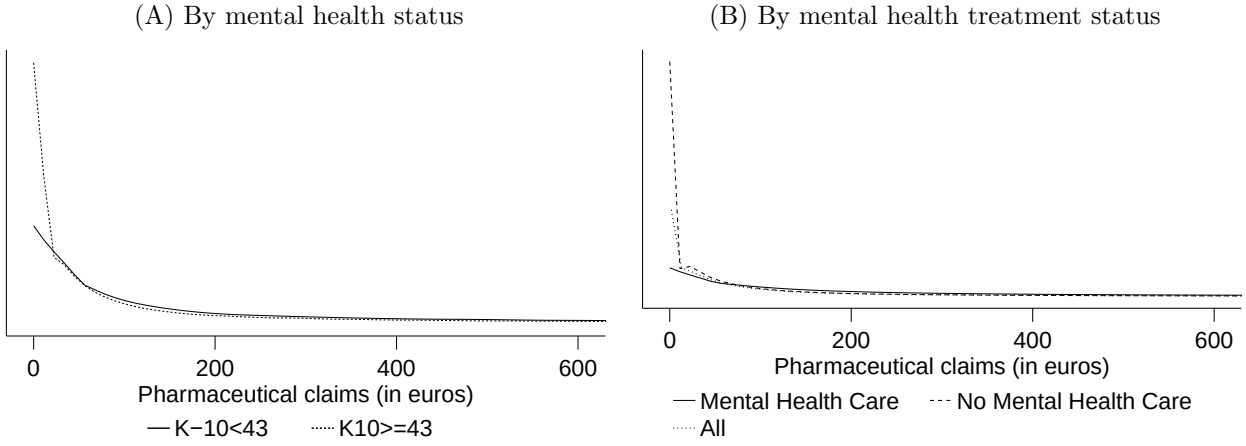

Figure B.10: Pharmaceutical claims in 2011

Notes: The figures show smoothed probability density functions. K-10 scores were measured in the 2012 GEMON survey, treatment status was assessed in 2011 for the full population.

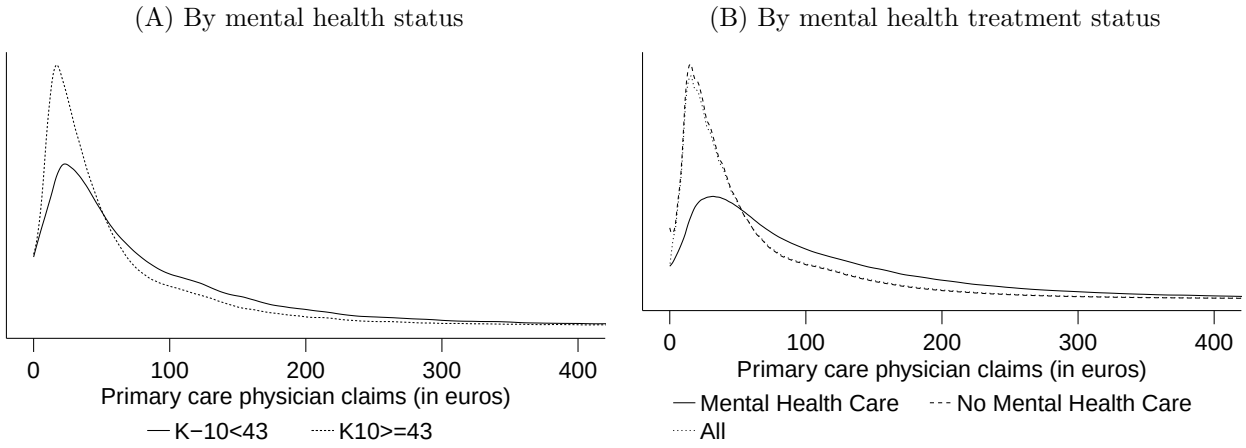

Figure B.11: PCP claims minus subscription fees in 2011

Notes: The figures show smoothed probability density functions. K-10 scores were measured in the 2012 GEMON survey, treatment status was assessed in 2011 for the full population.

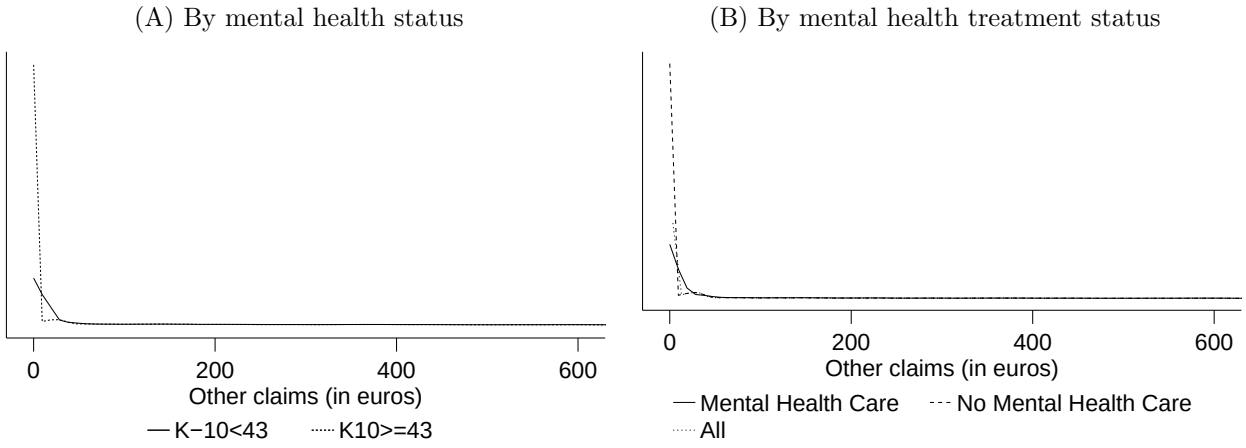

Figure B.12: Other health care claims in 2011

Notes: The figures show smoothed probability density functions. K-10 scores were measured in the 2012 GEMON survey, treatment status was assessed in 2011 for the full population.

	Adjustment disorder	Anxiety disorder	Bipolar disorder	Dementia	Depressive disorder	Disorder first diagn. in childhood
	(1)	(2)	(3)	(4)	(5)	(6)
Number of treatment records	88,270	146,291	44,195	8,292	210,209	82,386
Panel A: Characteristics (mean/median)						
Begin GAF	6.18	5.94	6.04	5.31	5.88	5.86
	(1.06)	(1.05)	(1.56)	(1.52)	(1.12)	(1.24)
Change in GAF	0.51	0.43	0.16	0.11	0.45	0.23
	(1.13)	(1.05)	(0.94)	(1.09)	(1.06)	(0.96)
Duration of DBC	229.51	283.41	322.00	246.42	285.11	293.62
	(125.27)	(111.59)	(97.25)	(138.50)	(113.07)	(107.12)
Total minutes	610	900	730	866	845	848
	[300–1,122]	[415–1,661]	[278–1,844]	[365–1,972]	[365–1,621]	[345–1,625]
Total visits	15.62	22.67	29.21	27.17	23.44	22.45
	(18.92)	(24.91)	(37.11)	(40.38)	(27.84)	(25.97)
Any inpatient days	0.04	0.04	0.14	0.16	0.07	0.04
	(0.20)	(0.20)	(0.35)	(0.37)	(0.25)	(0.18)
Total inpatient days if at least one	18	28	33	57	32	34
	[5–45]	[9–62]	[14–71]	[17–151]	[12–65]	[11–97]
Prescribed medication	0.19	0.37	0.72	0.32	0.47	0.50
	(0.39)	(0.48)	(0.45)	(0.47)	(0.50)	(0.50)
Panel B: Closing reason (percentage)						
Patient decision	8.69	8.1	3.57	5.28	7.42	7.25
Therapist decision	5.74	4.52	2.47	5.66	4.6	4.16
Mutual agreement	49.81	35.41	14.25	35.52	33.31	30.77
Open follow-up DBC	32.26	50.13	78.9	48.43	53.01	55.79
(Pre)intake for crisis	3.49	1.84	0.8	5.11	1.66	2.02

Table B.I: Treatment record characteristics by diagnosis

	Miscellaneous disorder	Missing	Other conditions	Personality disoder	Psychotic disorder	Substance-related disorder
	(1)	(2)	(3)	(4)	(5)	(6)
Number of treatment records	78,934	66,731	95,932	120,751	94,713	102,142
Panel A: Characteristics (mean/median)						
Begin GAF	5.99	5.81	6.23	5.76	5.08	5.37
	(1.15)	(2.04)	(1.23)	(1.16)	(1.42)	(1.08)
Change in GAF	0.43	0.18	0.43	0.22	0.10	-0.01
	(1.14)	(1.09)	(1.13)	(0.95)	(0.81)	(1.35)
Duration of DBC	264.10	80.28	253.34	299.88	320.53	268.55
	(118.02)	(92.27)	(119.12)	(107.01)	(101.21)	(122.33)
Total minutes	857	180	630	1,225	1,595	1,180
	[390–1,682]	[90–348]	[310–1,164]	[551–2,450]	[649–3,615]	[506–2,662]
Total visits	23.17	3.79	15.53	33.67	49.40	43.29
	(27.44)	(5.32)	(17.13)	(37.30)	(51.24)	(48.76)
Any inpatient days	0.06	0.01	0.02	0.09	0.21	0.24
	(0.23)	(0.11)	(0.13)	(0.29)	(0.41)	(0.43)
Total inpatient days if at least one	35	4	24	29	40	28
	[12–76]	[2–8]	[8–57]	[8–79]	[13–101]	[9–69]
Prescribed medication	0.22	0.01	0.10	0.38	0.67	0.41
	(0.41)	(0.08)	(0.30)	(0.49)	(0.47)	(0.49)
Panel B: Closing reason (percentage)						
Patient decision	9.9	0.4	10.46	7.69	3.32	12.16
Therapist decision	4.75	0.2	4.93	4.5	3.29	8.27
Mutual agreement	41.45	1.82	46.73	24.85	10.72	22.67
Open follow-up DBC	40.38	1.17	34.67	61.24	81.73	54.64
(Pre)intake for crisis	3.51	96.4	3.21	1.72	0.93	2.26

Table B.II: Treatment record characteristics by diagnosis (continued)

Table B.III: Average mental health score by treatment status in 2011

	No MHC (1)	MHC patients (2)
K-10 score	44.61	36.97
	(5.89)	(9.37)
MHI-5 score in 2011	19.99	14.94
	(3.58)	(5.20)

Notes: K-10 was measured in the GEMON survey in 2012, MHI-5 was measured in the GECON survey in 2011. Mental health care (MHC) treatment status was assessed in 2011.

Table B.IV: Median income and wealth by mental health and treatment status

	K-10 score 2012		MH treatment 2011	
	K-10≥43 (1)	K-10<43 (2)	No MHC (3)	MHC patients (4)
Personal income 2011 (€)	30,193	21,846	26,478	18,968
	[15,212–45,264]	[11,765–35,345]	[13,484–41,259]	[13,112–31,322]
Household income 2011 (€)	40,526	34,144	37,569	27,792
	[29,347–54,245]	[22,748–47,459]	[25,536–51,560]	[16,712–41,957]
Household wealth 2011 (€)	114,080	41,964	48,232	5,671
	[10,995–272,435]	[1,655–197,282]	[1,447–201,247]	[87–94,975]

Notes: Interquartile ranges between the 25th and the 75th percentiles in square brackets. The bottom two rows exclude youths under 18 years old in columns 1 and 2, but include youths in columns 3 and 4.

Table B.V: Mental health status and treatment by personal income

	% of pop (1)	K-10 score (2)	Proportion in MH treatment (3)
First positive income group	20.32	**42.38**	**0.116**
		(7.64)	
Second positive income group	20.32	0.86***	−0.035***
		(0.04)	(0.000)
Third positive income group	20.32	2.17***	−0.070***
		(0.05)	(0.000)
Fourth positive income group	20.32	3.24***	−0.084***
		(0.05)	(0.000)
Negative income	0.50	1.29***	−0.069***
		(0.21)	(0.001)
Zero income	5.77	1.53***	−0.069***
		(0.06)	(0.000)
Student grant	7.57	0.49***	−0.038***
		(0.09)	(0.000)
Pension	4.89	2.22***	−0.054***
		(0.07)	(0.000)

Notes: Adults aged 18–64. Personal income and treatment status are measured in 2011, K-10 scores come from the 2012 GEMON survey. Column 1 shows the proportion of the population in each income category. The first row of each panel shows the mean K-10 score (column 2) or the percentage in treatment (3) for the baseline category (standard deviations in parentheses). All other rows show regression coefficients which indicate the differences between each respective category and the baseline category (standard errors in parentheses).

Table B.VI: Mental health status and treatment by household income and wealth

	% of pop (1)	K-10 score (2)	Proportion in MH treatment (3)
Panel A: Household income			
First positive income group	24.74	**42.01**	**0.121**
		(7.86)	
Second positive income group	24.74	2.04***	−0.060***
		(0.04)	(0.000)
Third positive income group	24.74	2.95***	−0.076***
		(0.04)	(0.000)
Fourth positive income group	24.74	3.53***	−0.084***
		(0.04)	(0.000)
Negative income	0.56	2.28***	−0.068***
		(0.19)	(0.001)
Zero income	0.47	0.85**	−0.074***
		(0.36)	(0.001)
Panel B: Household wealth			
First positive wealth group	20.67	**41.70**	**0.111**
		(8.02)	
Second positive wealth group	20.67	2.51***	−0.051***
		(0.05)	(0.000)
Third positive wealth group	20.67	3.18***	−0.063***
		(0.04)	(0.000)
Fourth positive wealth group	20.67	3.55***	−0.069***
		(0.04)	(0.000)
Negative wealth	13.57	2.64***	−0.054***
		(0.05)	(0.000)
Zero wealth	3.75	−2.63***	0.047***
		(0.10)	(0.000)

Notes: Adults aged 18–64. Income, wealth, and treatment status are measured in 2011, K-10 scores come from the 2012 GEMON survey. Column 1 shows the proportion of the population in each income or wealth category. The first row of each panel shows the mean K-10 score (column 2) or the percentage in treatment (3) for the baseline category (standard deviations in parentheses). All other rows show regression coefficients which indicate the differences between each respective category and the baseline category (standard errors in parentheses).

Table B.VII: Mental health status and treatment by age, gender, and ethnicity

	% of pop (1)	K-10 score (2)	Proportion in MH treatment (3)
Panel A: Age			
18–24	13.36	**43.29**	**0.072**
		(6.51)	
25–34	18.51	0.56***	0.009***
		(0.06)	(0.000)
35–44	22.97	0.81***	0.004***
		(0.05)	(0.000)
45–54	24.02	0.70***	−0.004***
		(0.05)	(0.000)
55–64	21.14	1.41***	−0.027***
		(0.05)	(0.000)
Panel B: Gender			
Female	50.01	**43.56**	**0.076**
		(6.55)	
Male	49.99	1.38***	−0.016***
		(0.03)	(0.000)
Panel C: Ethnicity			
Native	79.90	**44.53**	**0.063**
		(6.00)	
Other Western countries	6.16	−1.43***	0.012***
		(0.06)	(0.000)
Other non-Western countries	3.70	−3.46***	0.009***
		(0.09)	(0.000)
Indonesia	2.87	−0.30***	0.011***
		(0.08)	(0.000)
Suriname	1.94	−5.09***	0.047***
		(0.15)	(0.001)
Morocco	2.35	−6.63***	0.058***
		(0.13)	(0.001)
Turkey	2.25	−2.66***	0.030***
		(0.13)	(0.001)
Dutch Antilles and Aruba	0.82	−1.76***	0.024***
		(0.19)	(0.001)

Notes: Adults aged 18–64. K-10 scores come from the 2012 GEMON survey, all other variables are measured in 2011. Column 1 shows the proportion of the population in each category. The first row of each panel shows the mean K-10 score (column 2) or the percentage in treatment (3) for the baseline category (standard deviations in parentheses). All other rows show regression coefficients which indicate the differences between each respective category and the baseline category (standard errors in parentheses).

Table B.VIII: Mental health status and treatment among first- and second-generation migrants

	% of pop (1)	K-10 score (2)	Proportion in MH treatment (3)
Native	79.90	**44.53**	**0.063**
		(6.00)	
Other Western countries first gen	3.30	-2.16^{***}	0.006^{***}
		(0.09)	(0.000)
Other Western countries second gen	2.86	-0.84^{***}	0.018^{***}
		(0.08)	(0.000)
Other non-Western countries first gen	3.13	-3.87^{***}	0.009^{***}
		(0.10)	(0.000)
Other non-Western countries second gen	0.57	-1.32^{***}	0.010^{***}
		(0.22)	(0.001)
Indonesia first gen	0.62	-0.39^{**}	-0.002^{**}
		(0.16)	(0.001)
Indonesia second gen	2.25	-0.28^{***}	0.015^{***}
		(0.10)	(0.001)
Suriname first gen	1.37	-5.58^{***}	0.057^{***}
		(0.17)	(0.001)
Suriname second gen	0.57	-3.27^{***}	0.025^{***}
		(0.33)	(0.001)
Turkey first gen	1.61	-7.75^{***}	0.072^{***}
		(0.15)	(0.001)
Turkey second gen	0.75	-3.89^{***}	0.027^{***}
		(0.23)	(0.001)
Morocco first gen	1.48	-3.06^{***}	0.033^{***}
		(0.15)	(0.001)
Morocco second gen	0.77	-1.78^{***}	0.023^{***}
		(0.22)	(0.001)
Dutch Antilles and Aruba first gen	0.58	-1.79^{***}	0.023^{***}
		(0.23)	(0.001)
Dutch Antilles and Aruba second gen	0.24	-1.66^{***}	0.027^{***}
		(0.35)	(0.002)

Notes: Adults aged 18–64. K-10 scores come from the 2012 GEMON survey, all other variables are measured in 2011. Column 1 shows the proportion of the population in each category. The first row of each panel shows the mean K-10 score (column 2) or the percentage in treatment (3) for the baseline category (standard deviations in parentheses). All other rows show regression coefficients which indicate the differences between each respective category and the baseline category (standard errors in parentheses).

Table B.IX: Mental health status and treatment by educational attainment and household type

	% of pop (1)	K-10 score (2)	Proportion in MH treatment (3)
Panel A: Educational attainment			
Doctoral or equivalent level	0.09	−0.00	−0.005
		(0.57)	(0.004)
Master's or equivalent level	10.47	**45.04**	**0.059**
		(5.09)	
Bachelor's or equivalent level	20.16	−0.27***	0.002***
		(0.07)	(0.000)
Short-cycle tertiary education	0.71	−0.77***	0.010***
		(0.21)	(0.001)
Post-secondary non-tertiary education	1.20	−0.84***	0.015***
		(0.17)	(0.001)
Upper secondary education	44.10	−1.30***	0.025***
		(0.07)	(0.000)
Lower secondary education	15.84	−2.49***	0.060***
		(0.08)	(0.000)
Primary education	5.39	−4.07***	0.083***
		(0.11)	(0.001)
Pre-primary education and development	2.05	−6.02***	0.086***
		(0.17)	(0.001)
Panel B: Household type			
Single adult	16.56	**42.61**	**0.116**
		(7.42)	
Member of couple without children	27.30	1.91***	−0.062***
		(0.04)	(0.000)
Member of couple with children	39.11	2.37***	−0.074***
		(0.04)	(0.000)
Single parent	4.27	−0.45***	0.008***
		(0.08)	(0.000)
Living in institutional home	0.24	−4.07***	0.051***
		(0.54)	(0.002)
Other	1.49	0.30**	−0.041***
		(0.14)	(0.001)
Child living at home	11.03	1.36***	−0.042***
		(0.08)	(0.000)

Notes: Adults aged 18–64. K-10 scores come from the 2012 GEMON survey, all other variables are measured in 2011. Column 1 shows the proportion of the population in each category. The second row of Panel A and the first row of Panel B shows the mean K-10 score (column 2) or the percentage in treatment (3) for the baseline category (standard deviations in parentheses). All other rows show regression coefficients which indicate the differences between each respective category and the baseline category (standard errors in parentheses).

Table B.X: Cause of death by mental health (treatment) status

	K-10 score		MH treatment	
	K-10≥43 perc (1)	K-10<43 coeff (2)	No MHC perc (3)	MHC patients coeff (4)
Death by suicide	0.000081	0.000230*** (0.000057)	0.000200	0.001208*** (0.000021)
Death by poisoning	NA	0.000081*** (0.000025)	0.000012	0.000125*** (0.000006)
Death by homicide	NA	−0.000005 (0.000010)	0.000011	0.000036*** (0.000005)
Death by organic mental disorders	0.000000	0.000018* (0.000010)	0.000017	0.000125*** (0.000006)
Death by substance use disorders	NA	0.000062** (0.000025)	0.000040	0.000312*** (0.000010)
Death by other mental disorders	0.000000	0.000031** (0.000014)	0.000010	0.000133*** (0.000005)
Death by other cause	0.004390	0.005684*** (0.000359)	0.005283	0.006175*** (0.000093)

Notes: Numbers refer to proportions. K-10 scores are observed in 2012, treatment status is observed in 2011, deaths are observed in 2013 and 2014, Results from separate regression models conditional on gender and year-of-birth dummies. Columns 1 and 3 give means, standard deviations in parentheses. Columns 2 and 4 give regression coefficients, standard errors in parentheses. Only adults aged 18-64.

Table B.XI: Claims by mental health (treatment) status

	K-10 score		MH treatment	
	K-10≥43 mean	K-10<43 coeff	No MHC mean	MHC patients coeff
	(1)	(2)	(3)	(4)
Total claims (in euros)	1,349	1,265***	1,320	3,958***
	(3,186)	(18)	(3,179)	(4)
Specialist mental health care claims (in euros)	46	295***	10	2,389***
	(523)	(5)	(222)	(1)
Hospital claims (in euros)	792	515***	793	589***
	(2,237)	(12)	(2,249)	(3)
Pharmaceutical claims (in euros)	202	192***	204	347***
	(555)	(3)	(583)	(1)
Primary care physician claims (in euros)	61	27***	64	51***
	(65)	(0)	(70)	(0)
Other claims (in euros)	181	119***	182	124***
	(600)	(3)	(601)	(1)

Notes: Results from separate regression models conditional on gender and year-of-birth dummies. Columns 1 and 3 give means, standard deviations in parentheses. Columns 2 and 4 give regression coefficients, standard errors in parentheses.

Table B.XII: Distribution of health care claims and long-term care use by mental health and treatment status (%)

	K-10≥43 (1)	K-10<43 (2)	No MHC (3)	MHC patients (4)
Panel A: Total claims				
Zero claims	0.37	0.23	0.78	0.08
Below median if claims> 0	51.84	35.47	52.61	9.02
Above median if claims> 0	47.78	64.30	46.60	90.90
Panel B: Hospital claims				
Zero claims	41.26	30.14	43.20	26.69
Below median if claims> 0	30.40	30.64	28.73	32.08
Above median if claims> 0	28.33	39.22	28.06	41.22
Panel C: Pharma claims				
Zero claims	28.40	17.57	29.59	11.35
Below median if claims> 0	36.65	31.89	36.61	25.21
Above median if claims> 0	34.95	50.54	33.80	63.44
Panel D: PCP claims				
Zero claims	0.55	0.38	1.16	0.49
Below median if claims> 0	52.53	36.96	51.17	26.03
Above median if claims> 0	46.92	62.66	47.67	73.48
Panel E: Other claims				
Zero claims	67.98	57.11	67.01	53.84
Below median if claims> 0	15.68	16.83	16.80	18.83
Above median if claims> 0	16.33	26.05	16.17	27.31
Panel F: Long-term care 2011				
No LTC	99.86	99.44	99.25	96.21
Nursing	0.06	0.29	0.24	1.00
Disability	0.06	0.14	0.44	0.58
Mental health	0.02	0.13	0.07	2.21
Panel G: Long-term care in 2013 if no LTC in 2011				
No LTC	99.96	99.79	99.9	98.96
Nursing	0.02	0.12	0.05	0.21
Disability	0.01	0.06	0.04	0.18
Mental health	NA	NA	NA	NA

Notes: Each column in each Panel sums up to 100%. We calculated median claims based for all columns on the full sample, which slightly deviates from median claims for the individuals within the full sample who had a nonmissing K-10 score. Mental health scores were measured in the 2012 GEMON survey, while treatment status was assessed in 2011 for the full population.

Table B.XIII: Long-term care use by mental health (treatment) status

	K-10 score		MH treatment	
	K-10≥43 perc (1)	K-10<43 coeff (2)	No MHC perc (3)	MHC patients coeff (4)
Panel A: Long-term care 2011				
Nursing institution	0.0006	0.0024***	0.0024	0.0080***
		(0.0002)		(0.0001)
Disability institution	0.0006	0.0007***	0.0044	0.0014***
		(0.0001)		(0.0001)
Mental health institution	0.0002	0.0011***	0.0007	0.0216***
		(0.0001)		(0.0001)
Panel B: Long-term care in 2013 if no LTC in 2011				
Nursing institution	0.0002	0.0010***	0.0005	0.0018***
		(0.0001)		(0.0000)
Disability institution	0.0001	0.0004***	0.0004	0.0014***
		(0.0001)		(0.0000)
Mental health institution	0.0000	0.0004***	0.0001	0.0064***
		(0.0001)		(0.0000)

Notes: Results from separate regression models conditional on gender and year-of-birth dummies. Columns 1 and 3 give means, standard deviations in parentheses. Columns 2 and 4 give regression coefficients, standard errors in parentheses.