# Life after a mental health diagnosis

Sonia Bhalotra[1], Bernt Bratsberg[2], and Andreas Kotsadam[2]

[1]University of Warwick, CAGE, IFS
[2]Ragnar Frisch Centre for Economic Research

September 11, 2026

**Abstract**

Using register data on the population of Norwegian workers, we investigate how people's lives change in multiple domains after they receive a (first) mental health diagnosis. Tracking more than 1.2 million workers through ten years, we observe that a mental health diagnosis is associated with sharp declines in employment, earnings, income net of transfers, partnership stability and fertility, alongside a deterioration in physical health and life expectancy. We additionally identify a spillover to the mental health of children. The estimates are robust to a battery of alternative designs exploiting diagnosis-timing, and co-worker and sibling comparisons. The estimates are similar by sex, by baseline somatic health, and across individuals in jobs characterised by different levels of occupational stress, social-interaction intensity or pay structure. Our findings are hard to reconcile with the view that diagnosis triggers effective treatment. They point to persistent mental health illness among at least a quarter of those initially diagnosed, career scarring, adverse impacts on family life and somatic health, and the tendency for reinforcing feedback across these outcomes as joint drivers.

# 1 Introduction

The lifetime prevalence of depression for adults is 1 in 5 for men and 1 in 3 for women (GALLUP, 2023), and depression accounts for only a third of all mental health illness (MHI). MHI constitutes an increasing share of the global burden of disease, and is now widely recognized as among the most challenging social problems of our times. MHI causes suffering to individuals, and to the people around them. In addition, they impair functionality and productivity, accounting for around 40% of disability and absenteeism among working-age individuals in high-income countries (Layard, 2013; OECD, 2012). Between one-third and one-half of new disability insurance claims in the OECD are attributed to mental disorders.[1]

In this paper we investigate the impact of a (first) mental health diagnosis on the employment and earnings of individuals as they evolve over a decade. We additionally model their mental health care trajectories, and other life outcomes including their probability of being in a cohabiting partnership, having children, suffering physical somatic morbidities, and their survival probabilities. These additional outcomes are of interest in themselves and they contribute to illuminating the mechanisms driving labour market outcomes.

The analysis uses longitudinal individual data linked across several Norwegian registers that cover the full population. We identify individuals age 20-65 who have a first mental health diagnosis either in primary care (under a GP) or in secondary care (under a mental health specialist) between 2009 and 2013, and we follow 423,534 such individuals together with 840,891 matched comparison individuals from five years before diagnosis to ten years after, a panel of 18.2 million person-years. We estimate how outcome trajectories change after diagnosis for diagnosed individuals relative to comparable individuals who have not yet been diagnosed.

The specification is an event study on matched cells with base-year-by-event-time fixed effects, in which comparison individuals contribute person-years only until their own (later) first diagnosis. We corroborate the main estimates with the Callaway and Sant'Anna (2021) estimator, and with a battery of alternative designs including a reweighted population event study balancing treated and untreated on age, sex, and base year; controls for base firm X year fixed effects and base occupation X year fixed effects to allow for firm and occupation specific shocks; colleague fixed effects, comparing workers within the same firm; same-sex sibling fixed effects, absorbing shared family background; and a design using only future-treated individuals as controls.

The remaining threat to identification is that the event of diagnosis is potentially en-

[1] Depression can reduce productivity through reducing the time endowment (days available to work), causing anhedonia which reduces the motivation to invest (by lowering perceived returns to investment), and causing fatigue and lost sleep which similarly reduce investment (by increasing the cost of investment; see Bhalotra et al. (2025); Angelucci and Bennett (2026) ). See, for instance, de Quidt and Haushofer (2016); Abramson, Boerma and Tsyvinski (2024), who formalize these insights from psychiatry, and Ridley et al. (2020) who review the association of depression with poverty.

dogenous. Mental health may deteriorate before individuals see a mental health professional and receive a diagnosis, mental health may deteriorate because employment or earnings are deteriorating, or an outside event such as job loss or bereavement may lead to a decline in both mental health and earnings. The pretrends in the five years preceding diagnosis do not support these concerns. To allow that we are not powered to detect differences in group trends, we demonstrate robustness of our findings to matching treated and comparison individuals on their earnings decile in each of the five years before diagnosis.

The estimates reveal a persistent decline in employment and earnings following a first mental health diagnosis. Ten years out, the employment rate is 7.9 percentage points lower, earnings conditional on employment are 6.5 percent lower, and (unconditional) annual real earnings are 17 percent down on the treated group's pre-diagnosis earnings. Roughly half the eventual earnings loss accrues within two years of diagnosis, after which the gap widens slowly, showing no sign of reversal at any point in the ten year horizon.

To illuminate the underlying mechanisms and sharpen interpretation of our findings we investigated a number of additional outcomes. First, we look at the extent to which welfare receipt buffers incomes. We estimate that transfers offset just under a third of the average earnings decline. The offset is concentrated in the first five years: transfers rise steeply for five years and then flatten, while earnings continue to fall. These results suggest that it is unlikely that individuals strategically obtain a mental health diagnosis in order to claim welfare. These results also provide us with a measure of the fiscal burden of mental health illness.

Second, to proxy persistence in the underlying condition we track mental health care contact. About a quarter of the treated (i.e. individuals with an initial diagnosis) have a mental health diagnosis in any given follow-up year, a share that is essentially flat from two years after diagnosis onward, and about a third of these carry the exact same diagnosis code year after year. The persistent earnings decline thus coexists with persistent illness for a substantial fraction. Moreover, those individuals that exit mental health care do not converge back to the control group's earnings path.

Third, we consider whether the diagnosis triggers occupational downgrading. We find that the treated group drift down the occupational ladder (relative to individuals who have not yet had a mental health diagnosis). Having ranked roughly 450 occupations by average earnings, we find that the treated fall about five rank points behind the control group within ten years, and this gap is as large conditional on employment. The earnings penalty thus includes both employment-exit and a drop in the quality of jobs held by those who stay employed. Our results are consistent with reinforcing feedback loops running between mental health, job quality, and human capital (skill acquisition on the job) (Jolivet and Postel-Vinay, 2025; de Quidt and Haushofer, 2016; Abramson, Boerma and Tsyvinski, 2024).

In fact, there is likely to be mutual reinforcement over the lifecourse not only between mental health and work-related outcomes, but between mental health, work, family life and

somatic health. Alongside a deterioration in labour market outcomes, we identify deleterious impacts of a mental health diagnosis on family life and somatic health. Ten years after the diagnosis, treated individuals are 8.6 percentage points more likely to be single, 4.8 points more likely to be divorced or separated, and 1.6 points more likely to be childless.

We find an increase in the Charlson comorbidity index that increases continuously through the event window, accumulating to 0.008 at year 9. This index estimates a patient's one-year mortality risk as a function of medical conditions and age, with higher values indicate a heavier burden of chronic disease and a greater risk of death. In line with this, we find that a first MH diagnosis is associated with a 26% higher mortality risk over the following 14 years, with a divergence apparent from year 4 onwards. This effect is large but consistent with estimates from psychiatry (Plana-Ripoll et al., 2023; Berk et al., 2023). These results tell us that the persistent earnings decline coexists with a real (physical) health gradient rather than (volitional) labour market disengagement.

These findings constitute a potential mechanism for the labour market results, while also being a potential consequence of those results (Altonji et al., 2021). Our observation that there is substantial feedback across the different life outcomes we consider is substantiated by our finding a significant age gradient in the losses that follow a mental health diagnosis. We find that younger individuals accumulate larger losses.

The Norwegian data allow us to examine heterogeneity along a number of other dimensions.[2] We find a broadly similar pattern of results when we investigate heterogeneity by sex, education and father's socioeconomic rank. There are some interesting differences, for instance, less educated individuals are more likely to drop out of employment, while more educated individuals are more likely to remain employed but in lower-pay occupations.

We have very rich data on clinical severity and condition. A notable finding is that cases that are plausibly classifed as mild in clinical terms (and that dominate numerically) suffer more persistent losses. This is consistent with acute cases being shortlived and/or with acute cases triggering more intensive treatment.

We hypothesised that individuals with weaker baseline somatic health would suffer larger losses on account of a compounding of somatic and mental health problems or if somatic health promotes treatment responsiveness, but we found no evidence of this. We thought it plausible that baseline occupation characteristics such as task content, job stress (following Deming, 2017) and the degree of social interaction entailed on the job would impact the extent to which a mental health diagnosis disrupts the labour market trajectory. However we find no discernible heterogeneity by these occupational traits. Considering that pay is more likely to reflect productivity in the private sector and in firms that offer bonus pay, we investigated outcomes by these cuts but again we find no statistically meaningful differences.

The persistent non-recovery of employment and earnings paths (and the long term im-

[2]We discuss the employment and earnings results here, results for other outcomes will be detailed in the text of the paper.

pacts on family and health outcomes) are hard to reconcile with the view that diagnosis triggers effective treatment. In studies of treatment effectiveness in psychiatry and medicine, the outcome is typically symptom remission and the samples analysed are relatively small. The handful of studies in economics that analyse impacts of mental health treatment on labour market outcomes, using designs that determine access to pharmacotherapy, psychotherapy, or insurance coverage often report mixed or null effects on employment.[3] Even in settings where treatment does help, only part of the earnings loss is offset (Biasi, Dahl and Moser, 2025; Shapiro, 2022). Several mechanisms could account for the observed patterns: clinical chronicity and high relapse rates in depression (Cipriani et al., 2018)[4]; gradual absorption into long-term sickness absence and disability insurance (Bratsberg, Fevang and Røed, 2010; Fevang, Hardoy and Røed, 2017; Maestas, Mullen and Strand, 2013); employer discrimination against workers with a history of mental illness (Bjørnshagen, 2021; Ridley, 2026); a tendency for people with common mental health disorders to forgo treatment (Abramson, Boerma and Tsyvinski, 2024); and reinforcing feedback between mental health, job quality, and human capital (Jolivet and Postel-Vinay, 2025; de Quidt and Haushofer, 2016; Abramson, Boerma and Tsyvinski, 2024).

We conclude with a statement of our contribution. This appears to be the first attempt to estimate causal impacts of a mental health diagnosis with broad diagnostic coverage on a cluster of key life outcomes at population scale over a ten year horizon. A handful of studies look at impacts of either mental health or mental health treatment on labour market outcomes alone, often in a field trial with a small sample and short horizon follow up, or focusing on a particular condition or treatment.[5]

Jolivet and Postel-Vinay (2025) estimate a structural model on longitudinal UK survey data that jointly models the dynamic impacts of mental health and job stress content on labour market choices, allowing job loss to aggravate mental health illness, and mental health to amplify the adverse effects of job loss. This sort of mutual reinforcement rationalizes the persistence of both mental health illness and earnings decline in our findings. de Quidt and Haushofer (2016) and Abramson, Boerma and Tsyvinski (2024) develop economic theories of

[3] Biasi, Dahl and Moser (2025) demonstrate that innovation of a treatment for bipolar disorder has significant impacts on employment, but bipolar accounts for 3 to 4% of all MHI. The evidence for treatment of common mental health disorders shows positive short run impacts on work hours and absenteeism (Shapiro, 2022; Lund et al., 2024; Angelucci, Fábregas and Vazquez, 2026), but no impact of treatment on employment (Angelucci and Bennett, 2024; Currie and Zwiers, 2025; Laird and Nielsen, 2016). In discussing this, Bhalotra, Daysal and Trandafir (2025) argue that early intervention might hold the key- they demonstrate large impacts of antidepressant treatment in childhood on future economic outcomes including employment, earnings and welfare dependence.

[4] Individuals who have experienced a mental health disorder are more than 50 times more likely to later develop dysthymia (chronic depression) than those who have no history of MDD; World Mental Health Surveys.

[5] See footnote 2 where we discuss the handful of studies in economics. Outside economics, much of the existing literature is cross-sectional (Kessler et al., 2008; Marcotte and Wilcox-Gök, 2003) or descriptive register-based (Hakulinen et al., 2019, 2020; Plana-Ripoll et al., 2023), though see (Andersen et al., 2024) who look at labour market impacts of PSTD following a factory explosion in Denmark.

mental health grounded in the psychiatry literature in which individuals experiencing mental illness have pessimistic expectations, lower self-esteem, and lose time due to rumination. As a result, they work less, consume less, invest less in risky assets, and forgo treatment which in turn reinforces mental illness. These formulations contribute to understanding our results.

The paper proceeds as follows. Section 2.1 describes the Norwegian registers and Section 2.2 the sample. Section 3 presents the main event-study results, the pre-trend analysis, and robustness. Section 4 compares GP-diagnosed and specialist-diagnosed cases and splits primary-care cases by type of code. Section 5 reports heterogeneity by sex, education and family background, by baseline somatic morbidity, by occupational stress and social-interaction intensity, by sector and pay structure, and by age at diagnosis. Section 6 turns to the other outcomes: income components, family outcomes, physical health and mortality, mental health care contact after diagnosis, and occupational rank. Section 7 concludes.

# 2 Data and Descriptive Statistics

## 2.1 Health Registers

We use two administrative health registers. KUHR (Kontroll og Utbetaling av HelseRefusjoner) records primary-care diagnoses made by general practitioners, coded in ICPC-2 and available from 2006. NPR (Norsk Pasientregister) records specialist and hospital diagnoses, coded in ICD-10 and available from 2008. Between them they cover the universe of publicly funded mental health contacts in Norway; only purely private, out-of-pocket treatment falls outside them.

The event study keys on the year of a person's first mental health diagnosis in either register. The binding data limitation is left-truncation at the start of each register. In the first years of coverage a contact is recorded as a first diagnosis whenever no earlier contact exists in the data, including for people already in treatment, so first-diagnosis counts are mechanically inflated at the beginning of the panel. They decline steeply thereafter and stabilise from about 2013. This is why the estimation sample draws first diagnoses from 2009 onwards rather than from the first year each register is available.

Table A1 traces coverage over the period. The stock of individuals ever diagnosed rises mechanically as the registers age, from 552,000 in primary care in 2008 to 1.39 million in 2023, and from 98,000 in the specialist register to 620,000 over the same period. Annual first-time flows move the other way, falling from 84,000 in 2008 to 22,000 in 2023 and flattening out after about 2013. The underlying population aged 20–65 is close to constant across these years, so neither pattern reflects demographic change.

The share ever diagnosed has shifted up at every age between 2013 and 2023. General practitioners account for most of the level and specialists add a smaller layer on top of it, and prevalence peaks between ages 25 and 40. Primary-care contact has become more frequent

alongside this. Visit rates rose at all ages over the same decade, whether measured by the share of people with at least one, three, six or twelve general-practitioner visits in a year, and whether computed over everyone aged 20–65 or only over those employed in the previous year. The number of distinct ICPC-2 diagnoses recorded per person rose in parallel; we use that breadth measure below to construct the baseline general-health cut.

Mental health codes sit high in the overall diagnosis ranking rather than in a specialist tail. In 2013 the most common primary-care codes were A97 (no disease, 3.4%), L90 (knee complaints, 2.5%), L01 (neck, 2.4%), K86 (hypertension, 2.4%), L02 (back, 2.1%) and P76 (depression, 1.5%), so a P-chapter mental health code appears among the top ranks alongside the musculoskeletal complaints that dominate primary-care contact.

## 2.2 Sample Construction and Descriptive Statistics

**Estimation sample.** The current design replaces reweighting with explicit matching. Each treated individual is matched to up to two comparison individuals of the same base year, age, and sex, and the regression absorbs a fixed effect for the resulting matched cell rather than for the individual. The base year is the year before the first diagnosis: base years run 2008–2012 and first diagnoses 2009–2013, so the base year is event time $-1$ and is the omitted category throughout.[6] The event window runs from five years before diagnosis to ten years after, sixteen event years in total. Treated individuals are aged 20–60 in the base year (birth cohorts 1948–1992); both groups must be resident in Norway on 1 January of both the base year and the diagnosis year; and comparison individuals stop contributing person-years from their own first diagnosis onward. Of the treated, 98.5% are matched to two comparison individuals and 1.5% to one. Table 1 reports base-year balance between the two groups.

Matching on base year, age, and sex balances those three margins by construction but leaves treated and comparison individuals unbalanced on education and on the labour market: in the base year the treated are 1.1 percentage points less likely to be employed, earn 52,500 NOK less, and are half again as likely to hold only compulsory education. This residual imbalance is what motivates the earnings-matched variants below, and it is the source of the pre-trend discussed in Section 3.3.

Restricting to individuals observed in all sixteen event years leaves 354,055 treated and 483,940 comparison individuals. The two earnings-matched designs are one-to-one. Matching on the single base-year decile costs almost nothing in sample size (423,480 of 423,534 treated retain a match); requiring a decile match in each of the five pre-years costs 46% of the treated sample, and also selects a less attached group (base-year employment 0.807 rather than 0.858). The estimation panel for the main design contains 18.2 million person-years

[6] Every base-year quantity reported below, including means, employment rates and matching cells, is therefore measured one year before diagnosis rather than in the diagnosis year.

Table 1: Balance at the base year, treated and matched comparison individuals

| | Treated | Comparison | Difference |
|---|---|---|---|
| Individuals at base year | 423,534 | 840,891 | |
| Female | 0.554 | 0.551 | +0.003 |
| Age | 39.49 | 39.54 | −0.05 |
| Compulsory education only | 0.222 | 0.151 | +0.072 |
| Upper secondary | 0.420 | 0.419 | +0.001 |
| Tertiary | 0.357 | 0.430 | −0.073 |
| Employed | 0.858 | 0.869 | −0.011 |
| Log earnings \| employed | 13.092 | 13.159 | −0.067 |
| Real earnings (2022 NOK) | 478,840 | 531,389 | −52,548 |

Education is missing for a small number of individuals, so the three education rows are computed on 422,197 treated and 839,190 comparison individuals rather than the full counts in the first row. Log earnings are computed on those with positive earnings (363,318 and 730,775).

for earnings and employment and 15.5 million for log earnings conditional on employment, clustered on 423,534 matched cells (423,332 for log earnings, which drops cells with no employed member).

Two features of the window matter. First, the age bounds pin it down exactly: the panel runs over ages 20–70, and the balanced subsample, individuals observed in all sixteen event years, spans ages 24–59 at the base year, which is precisely $20+4$ and $70-11$. That is what a window running from four years before the base year to eleven years after it implies, and it confirms the event-time indexing independently of the do-file.

Second, earnings for 2023 are not in the data. Event year +10 for the 2012 base-year cohort falls in 2023, so that cell is missing and is absorbed by the base-year × event-time fixed effects. The year-+10 coefficients therefore rest on the 2008–2011 base-year cohorts, four of the five, while every earlier event year uses all five. The +10 coefficient is identified from within-cohort differences among the cohorts that have the cell, so the effective sample at the final event year is smaller than elsewhere in the window.[7]

# 3 Main Event-Study Results

## 3.1 Design

Treatment is the year of first MH diagnosis in either register. Treated individuals have this year in 2009–2013. The comparison group consists of individuals with no MH diagnosis as of the base year; crucially, comparison individuals contribute person-years *only until their own*

[7]Note that the balanced subsample retains 2012 base-year individuals even so, because the balance criterion counts person-years present in the panel rather than person-years with a non-missing outcome. The balanced-panel row of Table 4 is therefore balanced on presence, not on observed earnings.

*first MH diagnosis*, after which they are dropped from the estimation sample. Event time runs from $-5$ to $+10$ around diagnosis, with the base year ($e = -1$) as the omitted category. Three outcomes are tracked: real earnings in 1000 2022 NOK, log earnings conditional on positive earnings, and a positive-earnings indicator.

The main specification is an event-study regression on matched cells:

$$y_{it} = \gamma_{p(i)} + \lambda_{b(i),e(i,t)} + \delta D_i + \sum_{e=-5,\, e\neq -1}^{+10} \beta_e \, \mathbf{1}[e(i,t) = e] \times D_i + \varepsilon_{it},$$

where $\gamma_p$ is a fixed effect for the matched cell $p$ (one treated individual and up to two comparison individuals of the same base year, age, and sex), $\lambda_{b,e}$ are base-year $\times$ event-time fixed effects (absorbing cohort-specific calendar shocks at every event time), $D_i$ indicates ever-treated, and the omitted category is event year $-1$. Standard errors are clustered on the matched cell. In Stata: `reghdfe y ib(9).timex##i.treat, absorb(baseyr_time pairid) cl(pairid)`, estimated on the sample that drops future-treated person-years from their own treatment year onward.

The base-year-by-event-time fixed effects allow the event-time coefficients to vary flexibly across diagnosis cohorts, allowing for changes in, for instance, diagnosis protocols, treatment practices, social norms and labour markets.

The matched-cell fixed effect is a weaker control than an individual fixed effect, and the difference is worth being explicit about. An individual fixed effect absorbs every time-invariant difference between people, personality traits, attitudes to work, family background, and identifies $\beta_e$ from within-person changes. The matched-cell fixed effect absorbs only what treated and comparison individuals share within a cell, so $\beta_e$ is a difference-in-differences in *group means* rather than in individual trajectories. The two coincide when the panel is balanced; they can differ when the composition of either group changes across event time. This matters here: the balanced-panel variant (individuals observed in all sixteen event years) gives a year-+10 earnings effect of $-57{,}971$ NOK against $-82{,}353$ in the main design, a difference of nearly a third; Section 3.4 discusses why we read the balanced figure as a lower bound rather than as a correction. What the matching buys in exchange is transparency about the comparison being drawn and the ability to match on the pre-diagnosis earnings path directly, which is what disciplines the pre-trends in Section 3.3.

We think it plausible that individuals in Norway in the study period, 2006–2023, approach a doctor within a year, which is the event time unit.

Individuals will often experience a decline in mental health before they make contact with a health professional and receive a diagnosis. This does not violate the integrity of the post-diagnosis decline in earnings but it requires that we test for earnings decline in the years preceding the diagnosis.

A diagnosis will typically trigger treatment, at least the supply of treatment. Making

contact with a health professional also implies that the individual has taken the first step towards demanding treatment. Additional relevant impacts of diagnosis include access to sick leave and welfare, and potentially to workplace responses including occupational downgrading, delayed promotion or discrimination.

**Why drop the not-yet-treated once they are treated.** Over a ten-year follow-up the initially undiagnosed comparison group is heavily contaminated: 28.5% of the comparison individuals matched in the September panel are diagnosed at some point during the window. Keeping their post-diagnosis person-years in the counterfactual drags the comparison path down and attenuates the estimated treatment effect. In the 2012-cohort test sample, the year-+10 earnings coefficient is −87,444 NOK when future-treated person-years are kept versus −106,937 when they are dropped (demographic-weights specification), an attenuation of roughly a fifth. Our design handles this the way the modern staggered-adoption literature recommends (Callaway and Sant'Anna, 2021; Borusyak, Jaravel and Spiess, 2024; Sun and Abraham, 2021): not-yet-treated individuals are valid controls only up to their own treatment date. Section 3.4 reports the formal Callaway and Sant'Anna (2021) estimator as a robustness check; it is computationally too heavy (days of run time on roughly 40 million person-year observations) to serve as the main specification, and our fixed-effects implementation reproduces its estimates closely.

**Alternative designs.** We retain the earlier battery of designs as robustness checks: (1) the reweighted population event study balancing treated and controls on age, sex, and base year; (2) colleague fixed effects within workplace; (3) same-sex sibling fixed effects; (4) future-treated individuals as the only controls. Results from these designs (estimated on the 5-pre/9-post window) are summarized in Section 3.4.

## 3.2 Main Results

Figure 1 shows raw trajectories of the three outcomes for treated and comparison individuals over the sixteen-year event window. In levels, treated and comparison earnings grow roughly in parallel through the pre-period, with the treated some 50,000 NOK below; the comparison group's earnings continue to grow after event year 0 while treated earnings flatten and never rejoin. The employment panel shows both groups peaking around the base year and declining thereafter, with the treated declining distinctly faster: the treated employment rate falls from 0.86 at the base year to below 0.75 by year +10, while the comparison group is still above 0.83.

Figure 2 reports the event-study coefficients and Table 2 the underlying numbers. In the pre-period all three outcomes are small relative to their eventual size; Section 3.3 discusses the pre-trend that remains. They move little in the diagnosis year and then fall sharply.

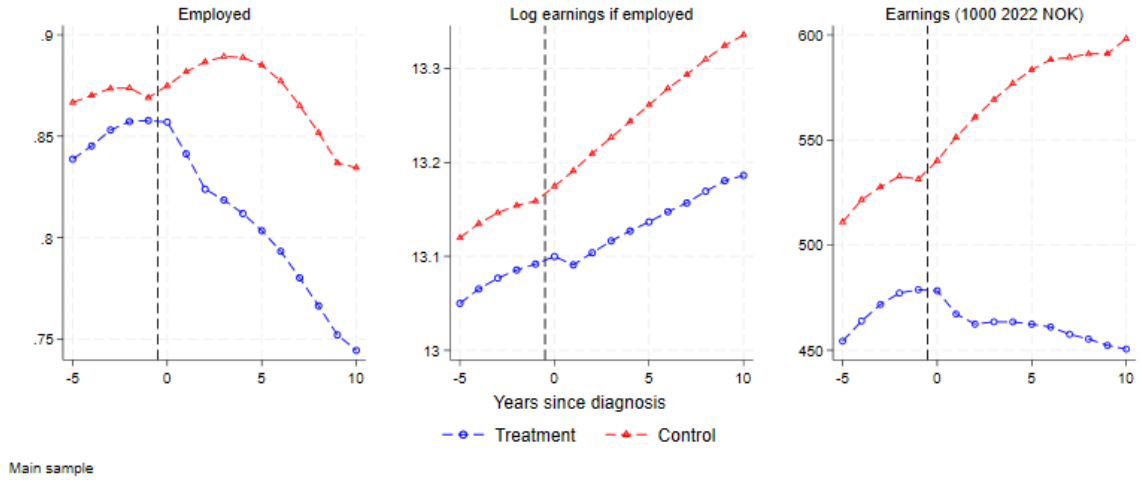



Figure 1: Earnings, log earnings conditional on employment, and employment by event time, treated vs. comparison group. Five years pre to ten years post diagnosis; base years 2008–2012, first diagnoses 2009–2013. Comparison individuals contribute until their own first MH diagnosis.

Employment is 0.65 percentage points below the base year in the diagnosis year, 2.9 points below one year later, 7.1 points below at +5, and 7.9 points below at +10 (s.e. 0.0009). Log earnings conditional on employment fall by 0.008 in the diagnosis year, 0.036 at +1, 0.059 at +5, and 0.065 at +10 (s.e. 0.0013). Real earnings fall by 9,085 NOK in the diagnosis year, 31,122 at +1, 44,978 at +2, 64,223 at +5, and 82,353 at +10 (s.e. 1,040).

The earnings decline at +10 is 17.2% of the treated group's base-year mean earnings of 478,840 NOK. Scaled against the earnings the treated would have had if they had grown at the comparison group's rate, the scaling used in the right-hand panel of Figure 2, the loss is about 15%. We report the NOK coefficient throughout and state the denominator whenever we give a percentage.

Three features of the path stand out. The break is sharp: two years after the diagnosis year the earnings gap is already more than half its eventual size (44,978 of 82,353 NOK). The decline then continues through +10 with no reversal. Real earnings fall at every event year; employment and log earnings flatten around +7 to +9, move back by less than 0.001, and widen again at +10. The extensive and intensive margins move together: the treated are both less likely to work and paid less when they do.

The path is not linear. About half of the eventual earnings loss accrues in the first two post-diagnosis years, and the annual increment falls from 22,037 NOK between the diagnosis year and +1 to 1,557 between +8 and +9. The gap stops widening quickly and then stays open.

Table 2: Main event-study estimates

| Event year | Employed | Log earnings \| employed | Real earnings (2022 NOK) |
|---|---|---|---|
| −5 | -0.0166 | -0.0074 | -3,950 |
| | (0.0007) | (0.0010) | (665) |
| −4 | -0.0136 | -0.0069 | -5,017 |
| | (0.0007) | (0.0009) | (843) |
| −3 | -0.0093 | -0.0058 | -3,315 |
| | (0.0006) | (0.0008) | (495) |
| −2 | -0.0053 | -0.0033 | -2,904 |
| | (0.0005) | (0.0007) | (652) |
| −1 | | *omitted (base year)* | |
| +0 | -0.0065 | -0.0078 | -9,085 |
| | (0.0005) | (0.0007) | (575) |
| +1 | -0.0290 | -0.0362 | -31,122 |
| | (0.0006) | (0.0008) | (849) |
| +2 | -0.0515 | -0.0435 | -44,978 |
| | (0.0006) | (0.0009) | (516) |
| +3 | -0.0597 | -0.0476 | -51,542 |
| | (0.0007) | (0.0009) | (553) |
| +4 | -0.0659 | -0.0530 | -58,147 |
| | (0.0007) | (0.0010) | (601) |
| +5 | -0.0705 | -0.0585 | -64,223 |
| | (0.0007) | (0.0010) | (636) |
| +6 | -0.0730 | -0.0615 | -68,583 |
| | (0.0007) | (0.0010) | (688) |
| +7 | -0.0739 | -0.0631 | -71,554 |
| | (0.0008) | (0.0011) | (740) |
| +8 | -0.0744 | -0.0629 | -73,733 |
| | (0.0008) | (0.0011) | (773) |
| +9 | -0.0740 | -0.0624 | -75,290 |
| | (0.0008) | (0.0012) | (858) |
| +10 | -0.0786 | -0.0651 | -82,353 |
| | (0.0009) | (0.0013) | (1,040) |
| Observations | 18,200,293 | 15,508,279 | 18,200,293 |
| Matched cells | 423,534 | 423,332 | 423,534 |

*Notes.* Coefficients on event-time × treated interactions from `reghdfe y ib(9).timex##i.treat, absorb(baseyr_time pairid) cl(pairid)`. Event year −1 (the base year, one year before diagnosis) is the omitted category. Standard errors clustered on the matched cell in parentheses. Log earnings are estimated conditional on positive earnings in the year concerned.

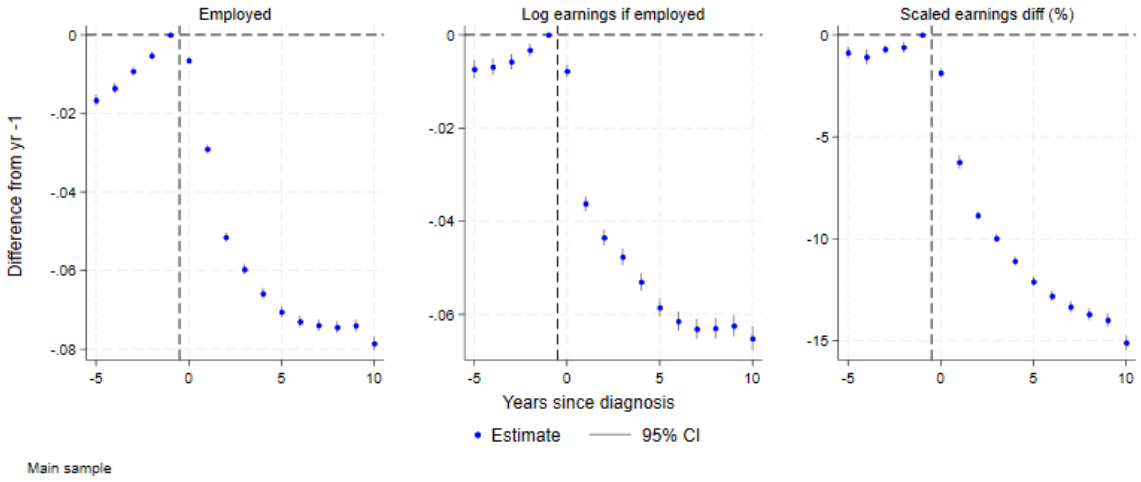



Figure 2: Main specification: event-study coefficients from matched cells (one treated individual and up to two comparison individuals matched on base year, age, and sex), with base-year × event-time fixed effects and future-treated person-years dropped from own treatment year onward. Difference from year −1 with 95% CIs. The right-hand panel scales the earnings difference by the treated group's base-year earnings grown forward at the comparison group's growth rate.

## 3.3 Pre-Trends and Anticipation

Matching on base year, age, and sex leaves treated and comparison individuals unbalanced on the labour market (Section 2.2), and the imbalance is not constant over the pre-period. In the full sample the employment lead is −1.66 percentage points at event year −5 and shrinks to −1.36, −0.93, and −0.53 at years −4, −3, and −2 before the normalisation at −1. Real earnings show a weaker and less regular version of the same convergence: −3,950 NOK at −5, then −5,017, −3,315, and −2,904 at years −4 to −2. Relative to their matched comparison individuals, the treated are catching up over the five years before diagnosis, and the normalisation at −1 is taken where the two groups are closest.

This pre-trend works against the finding. A treated group on an improving relative trajectory before the event would, if that trajectory continued, keep gaining on the comparison group after the event. Measured against that counterfactual, the post-diagnosis drop would be larger than the coefficients in Table 2. The pre-trend therefore does not produce a spurious effect, but the headline estimates should not rest on an assumption the data violate.

We therefore also match on the pre-diagnosis earnings path. The earnings-matched design pairs each treated individual one-to-one with a comparison individual in the same earnings decile in each of the five pre-diagnosis years, in addition to base year, age, and sex. Table 3 sets the two designs side by side and Figure 3 plots the earnings-matched estimates. Under the earnings match the leads are economically negligible: employment is within 0.13 of a

percentage point of the base year in every pre-year, and real earnings within 913 NOK against a +10 effect of 68,205 NOK. We do not describe the leads as zero, because with 229,530 matched cells the design can distinguish very small numbers from zero. The year −5 employment lead of −0.0012 has a $t$-statistic of −2.33 and is statistically significant while being two percent of the effect it precedes. The other eleven pre-period coefficients across the three outcomes are individually insignificant, with the largest $t$-statistic at −1.20.

Table 3: Pre-diagnosis leads under demographic and earnings matching

| | | Pre-diagnosis leads | | | | |
|---|---|---|---|---|---|---|
| Design | Outcome | −5 | −4 | −3 | −2 | +10 |
| Main (base year, age, sex) | Employed | -0.0166 | -0.0136 | -0.0093 | -0.0053 | -0.0786 |
| | Real earnings | -3,950 | -5,017 | -3,315 | -2,904 | -82,353 |
| Earnings-matched, $t-5$ to $t-1$ | Employed | -0.0012 | -0.0001 | -0.0001 | 0.0000 | -0.0635 |
| | Real earnings | 64 | -912 | -688 | -173 | -68,205 |

*Notes.* Event-time × treated coefficients, difference from event year −1. The upper design matches on base year, age, and sex (423,534 treated); the lower additionally matches on the earnings decile in each of event years −5 through −1 (229,530 treated, one-to-one).

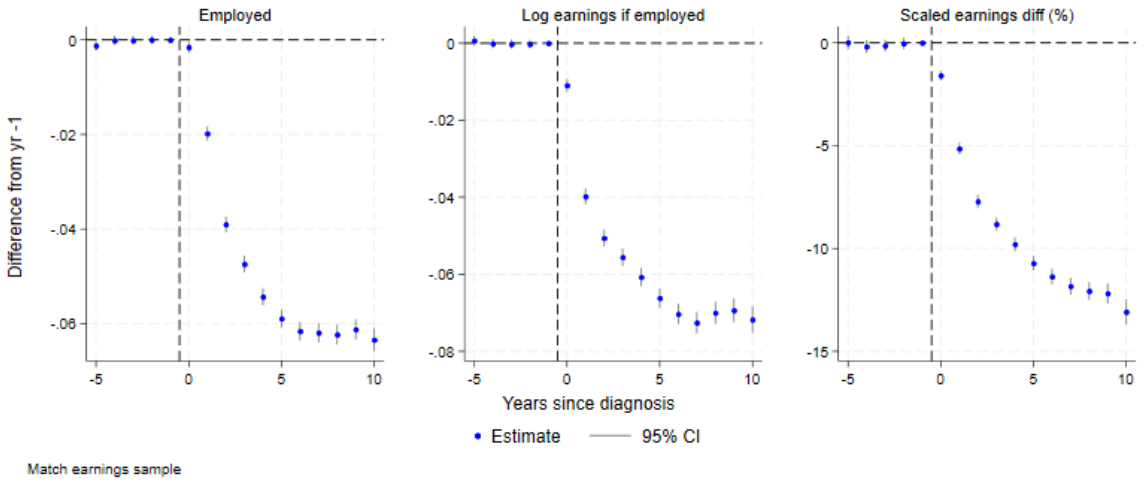



Figure 3: Earnings-matched design: each treated individual paired one-to-one with a comparison individual in the same earnings decile in each of event years −5 through −1, in addition to base year, age, and sex. Difference from year −1 with 95% CIs; 229,530 matched pairs. Compare the pre-period with Figure 2.

The result survives the stricter match. At event year +10 the earnings-matched design gives an employment effect of −6.35 percentage points against −7.86 in the main design, a log-earnings effect of −0.072 against −0.065, and an earnings effect of 68,205 NOK against

82,353. The gap between the two earnings coefficients is 14,148 NOK, a sixth of the headline. We do not attribute this sixth to the pre-existing gap alone, because the earnings match both tightens the comparison and drops 46% of the treated. Five-sixths of the headline earnings effect survives in a design where the pre-trend is economically negligible, and the intensive-margin effect is larger in that design.

Two caveats apply. Requiring a decile match in each of five pre-years costs 46% of the treated sample and selects a less attached group (base-year employment 0.807 against 0.858), so the comparison of magnitudes across the two rows of Table 3 combines the change in matching with the change in sample. And matching on five years of realised earnings is a strong conditioning set: if a deteriorating earnings path is an early symptom of the illness, matching on it removes part of the outcome. This is the usual trade-off in matching on pre-treatment outcomes, and it is why we report both designs. We keep the demographic match as the main specification, since every heterogeneity cut below is estimated on it, and use the earnings-matched design as the pre-trend check.

A variant matching on the earnings decile at $t-1$ only keeps almost the full sample (423,480 of 423,534 treated, one-to-one) but does not move the pre-trends. We therefore report the five-year version.

To see the shape of the pre-period beyond five years, we also estimate the matched design on a 21-year window with a symmetric ten-year pre-period. Over the longer pre-period the employment lead is about $-1.3$ percentage points at event year $-10$, falls to a trough of about $-1.75$ points around years $-8$ to $-6$, and then rises to the normalisation at $-1$. The scaled earnings difference has the same shape, starting slightly above zero at $-10$ and reaching about $-1.5\%$ at its lowest. The five-year window therefore begins just past the bottom of this trough, which is why its leads look monotone (Figure 4). The two runs agree where they overlap: the 21-year employment leads at years $-5$ to $-2$ are about $-1.7$, $-1.35$, $-0.9$, and $-0.5$ points, against $-1.66$, $-1.36$, $-0.93$, and $-0.53$ in the 16-year window, and the post-diagnosis paths are close (employment about $-7.7$ points at $+10$ against $-7.86$, and a scaled earnings loss of about 15% in both). The sample is identical in the two runs. Coefficients for the 21-year run are not yet tabulated; the figures in this paragraph are read from the plot.

The hump is informative. It is neither the steady pre-decline that gradually deteriorating health would produce nor a flat line. A dip six to eight years before diagnosis that has closed by the year before diagnosis is more consistent with an earlier episode of poor health, job loss, or family disruption that resolved than with the onset of the condition eventually diagnosed. We cannot separate these with the current data and leave it as an open question.

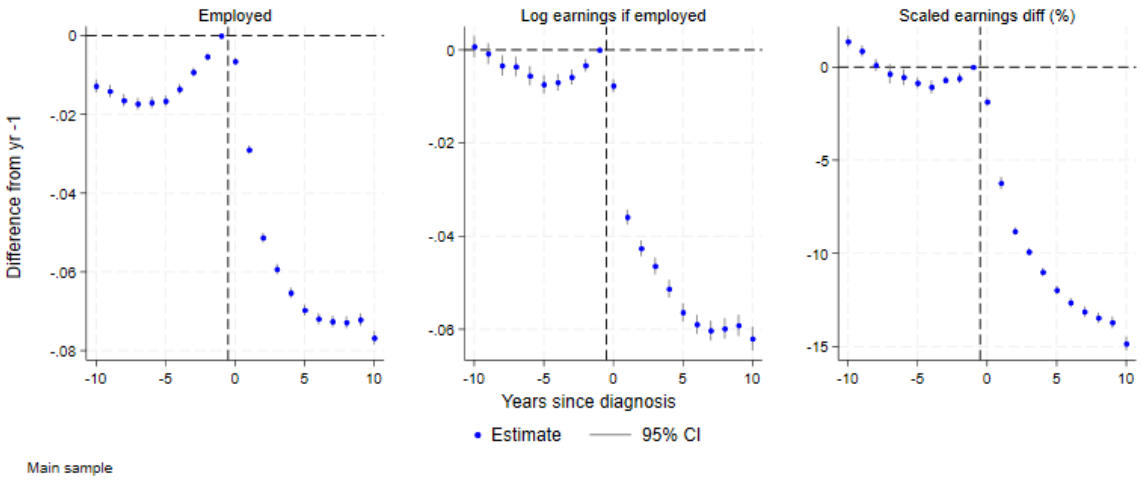



Figure 4: The same matched design estimated over a 21-year window with a symmetric ten-year pre-period, on the identical sample (423,534 treated). Difference from year −1 with 95% CIs. The pre-period is hump-shaped: near parity at −10, a trough around −8 to −6, and convergence back to the comparison group by −1. The sixteen-year window of Figure 2 begins on the recovering limb.

## 3.4 Robustness

**Sample and matching variants.** Table 4 collects the year +10 estimates across samples and matching rules. In the earnings column, the main design gives −82,353 NOK; restricting to individuals observed in all sixteen event years gives −57,971; restricting to those employed in the base year gives −75,470; dropping the 2008 and 2009 base years gives −83,726; and the earnings-matched design gives −68,205, or −55,329 when also balanced. Employment effects range from −5.4 to −7.9 percentage points and log-earnings effects from −0.044 to −0.072 across the six rows.

Dropping the two earliest base years changes almost nothing (1,373 NOK on 82,353), so the estimates are not driven by the earliest diagnosis cohorts. Conditioning on base-year employment moves the earnings coefficient by less than a tenth, so the result does not depend on who is working at the normalisation point.

The two restrictions that move the estimate are balancing the panel and matching on pre-diagnosis earnings. They move the earnings coefficient in the same direction for different reasons. Balancing the panel selects a more attached sample: base-year employment is 0.889 among the balanced treated against 0.858 overall, because requiring sixteen years of observation excludes people who leave the register, and post-diagnosis exit is itself an outcome. The −57,971 is therefore a lower bound. Matching on pre-diagnosis earnings selects a less attached sample (base-year employment 0.807), so the earnings-matched coefficient is not the main estimate purged of weak attachment; its effect comes from removing the level

gap between treated and comparison individuals. The two restrictions bracket the main estimate from different directions.

Table 4: Year-+10 estimates across samples and matching rules

| Design | Employed | Log earnings \| employed | Real earnings (2022 NOK |
|---|---|---|---|
| Main (matched on base year, age, sex) | -0.0786 | -0.0651 | -82,353 |
| | (0.0009) | (0.0013) | (1,040) |
| Balanced panel (all 16 event years) | -0.0618 | -0.0435 | -57,971 |
| | (0.0009) | (0.0013) | (1,126) |
| Employed in base year | -0.0716 | -0.0587 | -75,470 |
| | (0.0008) | (0.0013) | (1,135) |
| Base years 2010–2012 | -0.0743 | -0.0644 | -83,726 |
| | (0.0013) | (0.0019) | (1,767) |
| Earnings-matched, $t-5$ to $t-1$ | -0.0635 | -0.0718 | -68,205 |
| | (0.0012) | (0.0018) | (1,661) |
| Earnings-matched, balanced panel | -0.0536 | -0.0630 | -55,329 |
| | (0.0014) | (0.0019) | (1,947) |

*Notes.* Event-time × treated coefficient at event year +10, difference from event year −1. Rows 1–4 match on base year, age, and sex; rows 5–6 additionally match on the earnings decile in each of event years −5 through −1. Standard errors clustered on the matched cell.

**Callaway–Sant'Anna estimator.** The formal Callaway and Sant'Anna (2021) estimator, using not-yet-treated and never-treated individuals as the comparison group, was run on a test sample (the 2012 base-year cohort with its weighted control extract; 3.66 million person-year observations). Table A2 reports the dynamic ATT estimates for real earnings. The aggregate post-treatment ATT is −63,570 NOK (s.e. 1,347); the effect grows monotonically from −9,667 in the diagnosis year to −91,520 at event year +10. This tracks the main specification closely (test-cohort main spec at +10: −104,366; the CS estimate is somewhat smaller, consistent with its internal pair-balancing). Because a full-sample run takes days on roughly 40 million person-year observations, the CS estimator remains an appendix check rather than the main specification.

**The future-treated contamination, quantified.** The test cohort allows a clean comparison of how the treatment of comparison-group members over the follow-up period biases the naive design. On the 2012 cohort the year-+10 earnings coefficient under demographic weights is −87,444 NOK when future-treated comparison years are kept and −106,937 when they are dropped; under individual fixed effects, −84,996 against −103,344; and under individual plus base-year × event-time fixed effects, −98,429 against −104,366, the last dropping

future-treated person-years from their own treatment year onward, as the main specification does, rather than dropping those individuals entirely.

Keeping contaminated comparison years attenuates the estimate by roughly a fifth. All variants agree on the qualitative pattern; the main specification sits at the centre of the dropped-contamination estimates.

**Full-sample robustness: sample restrictions and matched pairs.** Table A3 collects the year-+10 earnings effect from the main specification and five full-sample alternatives. Restricting to workers with labour-market attachment five years before the base year (employed, or with positive wage income) lowers the estimate to about −73 to −74k, and restricting to the balanced panel of individuals observed in all 21 event years gives −69k: each selects a more attached sample and trims the effect modestly. Two matched-pairs designs, which link each treated individual to demographically matched controls and absorb base-year × event-time fixed effects, bracket the main estimate: demographic matching alone gives −99k, while matching that additionally conditions on industry and occupation gives −86k, essentially identical to the main specification. The agreement of the industry-and-occupation matched-pairs design with the fixed-effects main specification is reassuring, since the two rest on different identifying assumptions.

**Alternative identification strategies (5-pre/9-post window).** The four designs from the earlier specification round give the same qualitative result: a persistent post-diagnosis decline in earnings, log earnings, and employment, roughly linear in event time, with no recovery. At year +9: the reweighted population, colleague-FE, and same-sex-sibling designs each give earnings $\sim -75$k NOK, log earnings $\sim -0.07$, employment $\sim -7$pp; the future-treated-only control design gives $\sim -55$k, $\sim -0.055$, $\sim -6$pp. Figure A1 shows the reweighted-population and future-treated designs. Stability across these designs rules out selection on family background, on workplace shocks, and on the timing of diagnosis as drivers of the main result. The somewhat smaller future-treated estimate is expected: future-treated controls are themselves on a mild pre-diagnosis decline.

# 4 KUHR vs. NPR: Severity Selection

We split the treatment group by whether the first MH diagnosis comes from KUHR (GP/primary care) or NPR (specialist/hospital). The two registers differ both in the coding system (ICPC-2 vs ICD-10) and in the population they see: GPs capture a wide caseload of mild and common MH conditions, while specialists capture a smaller, more selected group at the severe end.

## 4.1 Estimates by Register of First Diagnosis

In the current panel, 369,367 of the treated have their first diagnosis in KUHR and 54,167 in NPR, against 840,891 comparison individuals. The specialist group is thus 12.8% of the treated.

Figure 5 plots the two series. The two groups follow different paths, and the difference takes the full ten-year window to appear. Employment falls further and faster for the specialist group early on: $-2.1$ percentage points in the diagnosis year against $-0.4$ for the GP group, and $-8.3$ against $-4.7$ at $+2$. It reaches $-8.4$ at $+3$ and then recovers, to $-7.5$ at $+5$, $-6.0$ at $+7$, and $-5.0$ at $+10$. The GP group's employment effect deepens throughout, from $-2.5$ points at $+1$ to $-8.3$ at $+10$. The two paths cross at about event year $+6$, and by the end of the window the GP group is the more affected by more than three percentage points.

Real earnings follow the same shape. The GP group's loss deepens to 86,278 NOK at $+10$, while the specialist group's reaches 62,900 at $+4$, flattens, and stands at 56,189 at $+10$. On the intensive margin the two separate further. For the GP group, log earnings conditional on employment fall throughout: $-0.033$ at $+1$, $-0.063$ at $+5$, $-0.074$ at $+10$ (s.e. 0.0013). For the specialist group they drop to $-0.062$ at $+1$ and then climb back: $-0.040$ at $+2$, $-0.024$ at $+5$, $-0.006$ at $+8$, and $+0.004$ at $+10$, where the estimate is indistinguishable from zero (s.e. 0.0036).

These coefficients are changes in a gap, not the gap itself. Specialist-diagnosed individuals earn substantially less than their matched comparison individuals before diagnosis: the level gap in log earnings at the base year is $-0.115$, against $-0.053$ for the GP group, and in real earnings 113,724 NOK against 43,219. Ten years after diagnosis the specialist group's shortfall in earnings conditional on employment is where it was the year before diagnosis ($-0.115 + 0.004 \approx -0.111$), while the GP group's has more than doubled ($-0.053 - 0.074 \approx -0.127$). Specialist-diagnosed individuals who stay employed do not close the gap; they stop losing ground, and the GP group does not.

Selection is part of this pattern. The specialist group loses employment sharply and does not fully regain it, so those who remain employed are an increasingly selected subset. A recovery in earnings conditional on employment is what that composition change would produce even with no improvement for any individual. The specialist pathway is associated with a large, early, partly reversible employment shock; the GP pathway with a smaller shock that does not reverse on either margin.

## 4.2 Symptom Codes, Disorder Codes, and Specialist Care

The register split can be taken one step further. Within KUHR, some first contacts are coded with a symptom-level ICPC-2 code that records a reason for encounter rather than a confirmed diagnosis (P29 "symptom not elsewhere classified", P06 sleep, P03 feeling de-

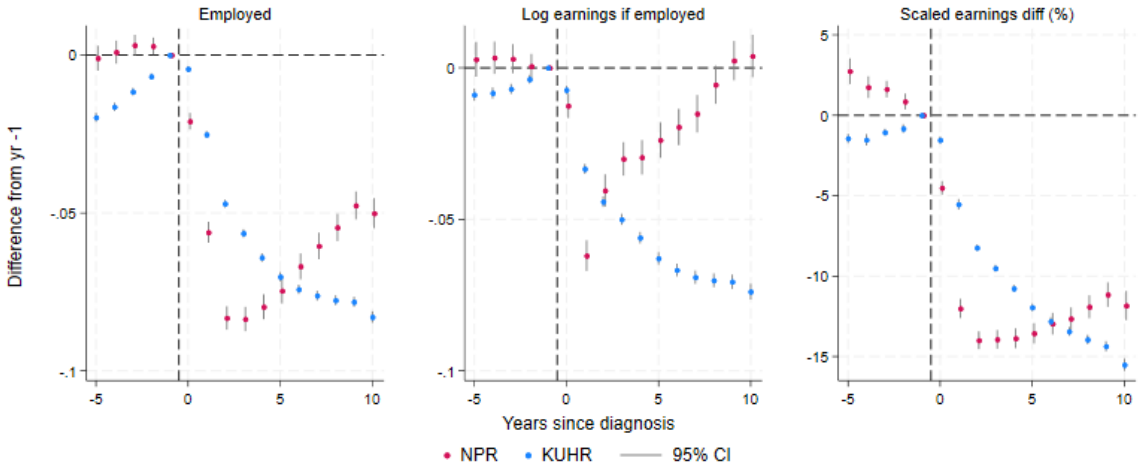



Figure 5: Event-study coefficients by register of first diagnosis. Difference from year −1 with 95% CIs. The specialist (NPR) group shows a deeper early employment loss that partially reverses, and a loss in earnings conditional on employment that returns to zero by year +10.

pressed, P01 feeling anxious), and others with a disorder-level code. Splitting the treated three ways gives 289,610 symptom-coded KUHR cases, 79,757 disorder-coded KUHR cases, and 54,167 specialist cases. The symptom-coded group is 68% of the treated.

If the coding hierarchy tracked severity, it would order the outcomes. Figure 6 shows that it orders them in reverse on all three margins. At event year +10 the symptom-coded group loses 8.60 percentage points of employment, the disorder-coded group 7.01, and the specialist group 4.99. Log earnings conditional on employment fall by 0.081 and 0.043 for the two primary-care groups, while the specialist group ends 0.004 above its baseline gap, indistinguishable from zero. Real earnings fall by 88,692, 76,100 and 56,046 NOK.[8] The mildest category by clinical reading is the most affected on every margin, and its intensive-margin path is the only one that continues to deteriorate: the disorder-coded group settles at about −0.04 from year +2, while the symptom-coded group falls from −0.033 at +1 to −0.081 at +10.

Two cautions apply. The symptom-coded group has the largest pre-trend of the three and is the only one entering the window below its comparison group on either margin: in log earnings at event year −5 it stands at −0.0123, against +0.0037 for the disorder group and +0.0027 for the specialist group, and in real earnings at −10,150 NOK against +4,361 and +11,028. Both overstate the contrast at the symptom end and understate it at the specialist end. And the specialist group is selected on staying employed, as Section 4.1 sets out. What

[8]The specialist coefficients differ trivially between this split and the register split of Section 4.1 (56,046 against 56,189 NOK) because they come from separate regressions with different comparison partitions, not because the group differs.

the split establishes is that the aggregate effect is not driven by the clinically severe minority. It is driven by the two-thirds of the treated whose first record is a reason-for-encounter code.

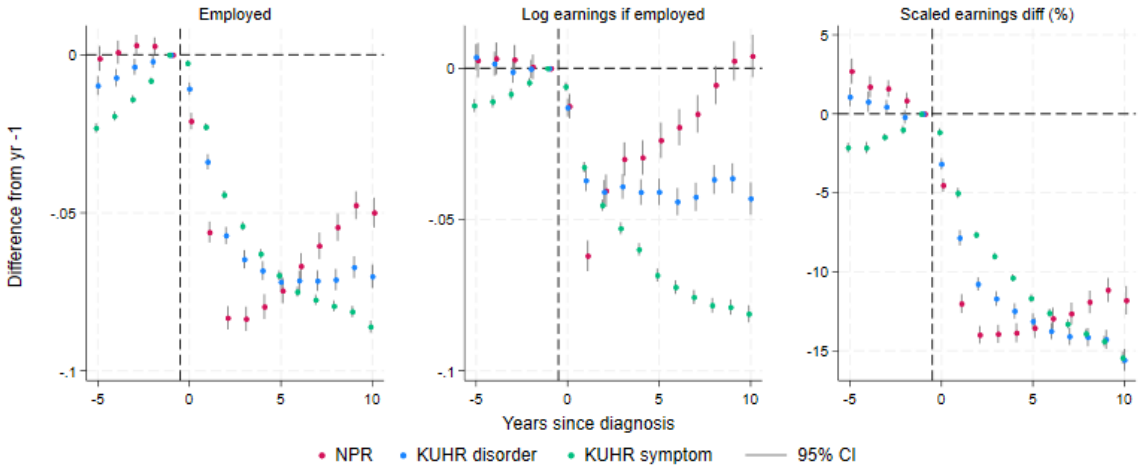



Figure 6: Event-study coefficients by symptom-coded KUHR, disorder-coded KUHR, and specialist (NPR) first diagnosis. Difference from year −1 with 95% CIs.

# 5 Heterogeneity

## 5.1 Sex, Education, and Family Background

We report three demographic cuts. Each is estimated on the main matched design by splitting the sample and re-estimating, so each row of Table 5 is a separate regression.

Father's social rank is constructed from fathers' labour earnings over ages 30–60, using all years for which earnings are observed (1967 and 2022 are the binding limits of the earnings register). Each father is ranked against all men in his own birth cohort, so the measure is a within-cohort percentile rather than an absolute income level. The sample is then split at the median father's rank within the estimation sample.

Figure 7 plots the education and father's-rank splits and Table 5 reports the estimates. Men lose 8.2 percentage points of employment at year +10, women 7.4; log earnings conditional on employment fall 0.069 for men and 0.067 for women; real earnings fall 100,041 NOK for men and 67,603 for women. Individuals without a college degree lose 7.7 points of employment against 6.7 for the college-educated, but 0.050 of log earnings against 0.073, and 57,690 NOK against 92,472. Individuals whose fathers are below the median in social rank lose 7.1 points of employment against 8.0 for those above, 0.047 of log earnings against 0.074, and 58,409 NOK against 92,255.

Table 5: Year-+10 estimates by sex, education, and father's social rank

| | Effect at +10 | | | Base-year | Loss as % | Earnings |
|---|---|---|---|---|---|---|
| Group | Employed | Log earn. | Real earn. | earnings | of base yr | lead at −5 |
| Men | -0.0817 | -0.0694 | -100,041 | 568,918 | -17.6 | 3,862 |
| | (0.0013) | (0.0020) | (1,946) | | | |
| Women | -0.0745 | -0.0667 | -67,603 | 406,453 | -16.6 | -10,275 |
| | (0.0012) | (0.0017) | (945) | | | |
| No college | -0.0768 | -0.0499 | -57,690 | 418,175 | -13.8 | -13,501 |
| | (0.0012) | (0.0017) | (1,026) | | | |
| College | -0.0666 | -0.0728 | -92,472 | 588,469 | -15.7 | 2,378 |
| | (0.0013) | (0.0020) | (2,085) | | | |
| Father below median rank | -0.0710 | -0.0475 | -58,409 | 443,122 | -13.2 | -4,717 |
| | (0.0014) | (0.0019) | (1,193) | | | |
| Father above median rank | -0.0798 | -0.0741 | -92,255 | 518,562 | -17.8 | 1,982 |
| | (0.0012) | (0.0018) | (1,645) | | | |

*Notes.* Event-time × treated coefficients at event year +10, difference from event year −1, from separate regressions on each subsample of the main matched design. Standard errors clustered on the matched cell in parentheses. "Base-year earnings" is the mean of real earnings among the treated in that subgroup at event year −1, and the "loss as %" column expresses the real-earnings coefficient as a percentage of it. The final column is the real-earnings coefficient at event year −5, included because it differs systematically across these subgroups and is what makes the ordering unreliable; see the discussion below.

Taken at face value, the education and family-background results show a gradient. Absolute earnings losses are about 60% larger for the college-educated than for those without a degree, and about 58% larger for those with high-ranked fathers than for those with low-ranked fathers. The advantaged groups start from higher base-year earnings, so absolute magnitudes overstate the gap, but the ordering survives normalisation: 15.7% of own base-year earnings against 13.8% by education, and 17.8% against 13.2% by father's rank.

**Pre-diagnosis trends across these groups.** The last column of Table 5 reports the real-earnings coefficient at event year −5, and it differs systematically across the groups. The disadvantaged groups enter the window well below their comparison individuals and converge upward; the advantaged groups enter slightly above parity and drift down to it. Because the normalisation is taken at −1, this compresses the measured loss for the groups that appear to lose least. We therefore report the ordering descriptively and do not build on it. Nor do we treat the earlier finding of no socioeconomic gradient as confirmed. The three cuts remain to be re-estimated on the earnings-matched design.

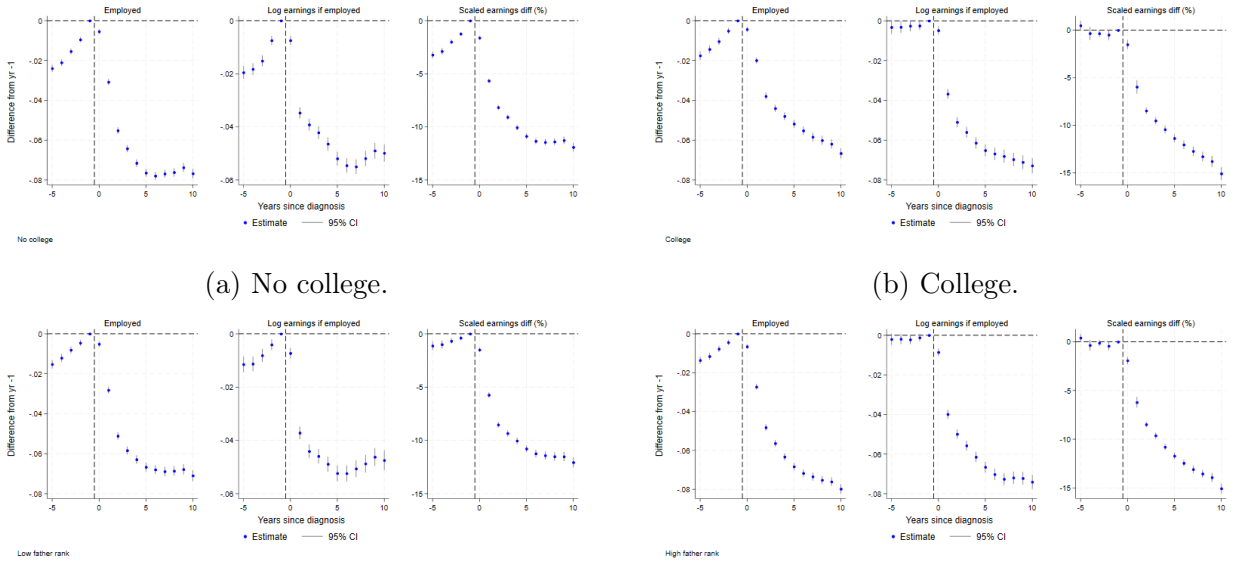
The composition of the loss also differs by education. The non-college group loses more employment and less earnings per worker; the college group loses less employment and more

earnings per worker. A diagnosis pushes lower-educated workers out of work and higher-educated workers down within work, into lower-paid roles, reduced hours, or interrupted careers, and over ten years the second carries the larger monetary cost.

If the gradient survives re-estimation, two readings are available. The first is affordability: welfare replacement in Norway is capped, so it falls furthest short of prior earnings for high earners, and the households best able to absorb a period out of work also have the most earnings to lose. On this reading the larger proportional loss among advantaged groups reflects capacity to step back rather than more severe illness. The second is that returns to continuous employment are steeper in the jobs the advantaged hold: careers with promotion ladders, firm-specific capital, and seniority-linked pay punish an interruption more than jobs where pay is flat in tenure.

Both readings weigh against a strategic-diagnosis interpretation, as do the transfer results in Section 6. If diagnoses were sought to access benefits, the incentive would be strongest where benefits replace the largest share of forgone earnings, at the bottom of the earnings distribution, and losses would be concentrated there. They are not.

The sex result does not move. Men's and women's trajectories after diagnosis are close on both margins, and the larger absolute earnings loss for men (100,041 against 67,603 NOK) is a level effect: 17.6% against 16.6% once each is scaled by its own base-year mean. Women are 55.4% of the treated but do not fare differently once diagnosed.




(a) No college.

(b) College.

(c) Father below median rank.

(d) Father above median rank.

Figure 7: Event-study coefficients by education and father's social rank. Difference from year −1 with 95% CIs.

## 5.2 Baseline Morbidity

We use two measures of baseline somatic morbidity: (i) the raw count of non-MH ICPC-2 diagnoses in primary care over a three-year lookback, split into terciles; and (ii) an extended Sandvik-based morbidity index over 24 somatic categories, coded as any versus none, which classifies 20% of the sample as non-healthy. P-chapter codes are excluded throughout to avoid mechanical confound with the treatment variable.

**Diagnosis-count terciles.** The terciles are uneven because the underlying count is discrete. The lowest tercile holds everyone with zero or one non-MH diagnosis in the three-year window (884,173 observations, 51% of the sample); the middle tercile is exactly two diagnoses (311,078, 18%); and the highest tercile is three or more, up to a maximum of 32 (535,332, 31%).

Treatment rate is markedly higher in T3 (34%) than in T1 (20%), confirming that baseline physical morbidity and MH diagnosis are positively correlated.

**Binary morbidity index.** 23.9% of treated are classified as non-healthy ($\geq 1$ morbidity category); 20.7% of controls are. The most common components for treated are musculoskeletal (11.9%), hypertension (4.6%), asthma (2.4%), thyroid (1.7%), migraine (1.7%), skin (1.2%), diabetes (1.4%).

**By diagnosis-count tercile.** Earnings losses at year +9, relative to year −1, are similar across the three terciles and, if anything, largest for the healthiest group: about −80k NOK in the lowest tercile, −65k in the middle and −70k in the highest. Log-earnings and employment effects are essentially identical across terciles ($\sim -0.07$ and $\sim -0.08$). The earnings-level heterogeneity is small and (weakly) in the opposite direction to the hypothesis.

**By the binary morbidity index.** The binary index gives the same picture: about −80k NOK for those with no recorded somatic morbidity at baseline and −60k for those with at least one. Again, log-earnings and employment effects are similar for the two groups ($\sim -0.07$ and $\sim -0.07$–$-0.08$). The level difference in earnings likely reflects higher baseline earnings among the healthy group (more to lose in absolute kroner), not a difference in relative treatment response.

The hypothesis was: MH treatment is *more effective* for individuals in *better* baseline general health, which would generate *smaller* post-diagnosis earnings drops for healthy individuals.

The data do not support this hypothesis, as Figure A2 shows. In kroner, healthier individuals lose weakly more, the opposite of the predicted direction, although the difference

is modest and could reflect differences in baseline earnings levels. In relative terms, in log earnings and in employment, there is no heterogeneity by baseline morbidity to speak of.

Two possible readings: (i) the level result is the real effect and baseline health buffers *against* the loss rather than facilitating treatment response; (ii) there is no meaningful treatment-effectiveness heterogeneity by baseline morbidity. Either way, the baseline-health mechanism from the earlier discussion receives little empirical support.

This is a substantive finding for the paper: the persistent earnings decline is not concentrated among the "unhealthier" treated group, which weakens one of the proposed mechanisms for why diagnosis fails to trigger effective treatment. Two supplementary checks would sharpen it: a Charlson index built from ICD-10 codes in the specialist register, and the alternative treatment-trajectory analysis.

## 5.3 Occupational Stress

We compute two indices. The first is narrow and uses only the three O*NET items that map most directly to psychological stress: stress tolerance, time pressure and consequence of error. It is measured at the ISCO-88 four-digit level and covers about 435 occupations. The second is broader: a 29-item principal-component stress index, measured at the ISCO-88 three-digit level and covering 111 occupations. Both are split into tertiles at the base year.

Both indices are split into tertiles. The two indices are strongly correlated but not identical: about 60% of person-years fall on the diagonal of the tertile cross-tabulation.

**Which occupations the indices pick out.** Under the narrow index, the top tercile contains teachers, nurses, electricians, chief executives, secretaries, truck drivers, bank clerks, engineering technicians and managers. The bottom tercile contains shop salespersons, who are the largest single group, together with childcare workers, public-administration clerks, cleaners, carpenters, pre-school teachers, caretakers and waiters. The broader index sorts occupations similarly at both ends: teachers, public administration, nurses, technical sales, chief executives and pre-school teachers at the top, and shop sales, cleaners, carpenters, truck drivers, engineering technicians, auto mechanics and waiters at the bottom.

Some intuitively "stressful" occupations, such as truck drivers and engineering technicians, sit in the low-stress tercile, because the O*NET items weight cognitive and emotional complexity more heavily than physical demands or time pressure. Both indices are therefore better read as measures of psychological job demands than of subjective stress.

**Psychological-stress index.** Earnings losses at year +9 are flat across the tertiles (Figure A3): about −70k NOK in high-stress occupations and −75k in both the medium- and the low-stress group. Log-earnings and employment effects are effectively identical across tertiles ($\sim -0.07$ and $\sim -0.07$).

**Broad job-stress index.** The broader index gives the same pattern: about −70k NOK for high stress, −75k for medium and −70k for low. Log-earnings: high stress ∼ −0.06, medium ∼ −0.075, low ∼ −0.07. Employment: high ∼ −0.06, medium ∼ −0.085, low ∼ −0.08.

We also estimate within-workplace versions of both indices, adding workplace fixed effects. The patterns are indistinguishable from those in the full sample.

The heterogeneity by occupational stress is essentially flat. In relative terms (logs, employment), neither index detects meaningful differences across stress tertiles. The high-stress tercile shows a slightly smaller drop than the medium- and low-stress tertiles under the broader index, but this is not a precise or robust finding.

This is not what one might expect under a pure "job demands" story where MH conditions most disrupt cognitively demanding work. Three readings are possible. High-stress occupations are also high-paying and high-education, with more scheduling flexibility and better access to treatment, and these may offset the demand-side vulnerability. The persistent earnings decline may run not through occupation-specific mechanisms but through exit from employment across occupations. Or the O*NET indices may fail to capture the subjective dimension of stress that matters for mental health.

This finding limits the "job demands" mechanism as a driver of the main result.

## 5.4 Social-Interaction Intensity

We compute two composite indices at the ISCO-88 four-digit level, covering 435 occupations, and split each into tertiles. The first combines seven items: three from Work Context (contact with others, face-to-face discussions, and dealing with external customers or the public) and four from Work Activities (establishing interpersonal relationships, assisting and caring for others, performing for or working with the public, and selling or influencing others). The second is a six-item alternative that draws only on sub-elements of communicating and interacting: interpreting information, communicating with supervisors, peers and subordinates, communicating with people outside the organisation, establishing interpersonal relationships, resolving conflicts and negotiating, and performing for or working with the public.

The two indices agree on tertile placement for ∼80% of person-years; the broad index pulls more occupations into the middle tertile.

All regressions here include workplace fixed effects and firm-balance weights, so the comparison is within firm.

**Tertile composition.** The tertiles are close to equal in size: 516,783 observations (33%) in the low tertile, 532,238 (34%) in the middle and 501,624 (32%) in the high tertile. Treatment rate is similar across tertiles (∼24–29%), with a slight tilt toward T2.

**Occupations by tertile.** The tertiles separate occupations as one would expect. The high-interaction tertile contains shop sales workers, teachers, care workers, nurses, sales representatives, chief executives and pre-school teachers. The low-interaction tertile contains cleaners, engineers, electricians, public-administration clerks, carpenters, engineering technicians and caretakers.

**Narrow index.** At year +9 the earnings losses are about −85k NOK in the high-interaction tertile, −70k in the middle and −75k in the low tertile (Figure A4). Log-earnings and employment effects are essentially identical across tertiles ($\sim -0.07$ and $\sim -0.07$). 95% confidence intervals overlap throughout.

**Broad index.** Pre-period level differences are sharper (high-social earn ∼200k more), but post-diagnosis trajectories are again very similar across tertiles (Figure A5). Earnings drop at year +9: high $\sim$ −90k, medium and low $\sim$ −75k. CIs overlap.

The within-firm event-study effects are essentially flat across social-interaction tertiles in both indices. This is informative as a null: depression, anxiety and stress diagnoses do not appear to "hit harder" in jobs that demand more interpersonal contact, at least under these indices and within firms.

Three caveats apply. Within-firm comparisons absorb the variation an across-firm cut would pick up, including firm-specific job design, accommodation policy and sorting. The O*NET items describe requirements at the level of the occupation and not job-design differences within one. And high-interaction occupations may also offer more job security or accommodation, which would offset the demand-side vulnerability.

This finding, together with the flat heterogeneity by occupational stress (Section 5.3), suggests that the persistent earnings decline is not strongly driven by occupation-specific mechanisms.

## 5.5 Sector and Pay Structure

We split the sample from an earlier estimation round (base years 2009–2013, nine-year follow-up) into three categories: public-sector employees, private-sector employees in occupations with bonus intensity below the within-private median, and private-sector employees in occupations above the median. Bonus intensity is measured as either the share of workers in the occupation who receive a bonus or the share of total wage income paid through bonuses.

**Sample composition at base year.** The three categories are similar in size. The public sector has 560,330 workers, 24% of them treated, with mean personal and firm bonus shares of 1.1% and 0.3%. Private-sector workers in occupations below the median in bonus intensity

number 553,226, 21% treated, with bonus shares of 3.6% and 0.9%. Those above the median number 555,663, 20% treated, with bonus shares of 10.0% and 2.3%.

MH prevalence is slightly higher in the public sector (24%) than in either private category (20–21%).

**Event-study results.** At year +9, plotted in Figure A6, public-sector workers lose about −55k NOK in earnings, 0.05 log points conditional on employment, and 7 percentage points of employment. In the private sector the losses are larger and almost identical in the low- and high-bonus groups: about −85k NOK, 0.07 to 0.075 log points, and 8 percentage points.

The public sector is meaningfully more protective on earnings and log earnings; the employment-margin difference is small. Within the private sector, bonus intensity does *not* matter: the two private series sit on top of each other at every event-time lag. Reading: the public-private gap reflects wage rigidity, tenure protection, or accommodation, not bonus exposure.

## 5.6 Age at Diagnosis

We split the treated and matched-control samples by age at diagnosis: below 50 ($N$ = 1,141,094) versus 50 and above ($N$ = 528,707). The under-50 group accounts for about 70% of treated cases; mean age at base year is 33.9 in the young group and 57.3 in the old group.

**Event-study results.** At year +9, plotted in Figure A7, those diagnosed before 50 lose about −95k NOK in earnings, 0.07 log points and 9 percentage points of employment, against −35k NOK, 0.05 log points and 6 percentage points for those diagnosed at 50 or later.

Young workers carry the bulk of the labor-market response. Older controls are mechanically leaving employment for retirement, which compresses the level differences; the absolute earnings drop for older treated is small because they sit on lower earnings trajectories already.

**Top first-diagnosis codes by age group.** The diagnosis mix differs by age. Among under-50 KUHR-first cases, the top codes are P02 acute stress (28%), P29 other psych symptom (17%), P76 depression (14%), and P06 sleep disturbance (13%). Among ≥ 50 KUHR-first cases, sleep (P06) moves up to the second-most-common code (20%) and depression (P76) sits at 12%; the share of acute-stress codes (P02) falls to 24%. NPR-first cases show the same pattern: F32 depressive episode and F43 stress dominate at all ages, but F10 alcohol disorders rise from 8% in the young to 12% in the old. The shift toward sleep complaints with age reinforces the case for a dedicated P06 split.

# 6 Additional Outcomes

This section collects the remaining outcomes, all estimated on the main event-study setup and the same full sample as Section 3.

Table 6 reports six further outcomes on the current panel.

Table 6: Income and family outcomes

| Outcome | +0 | +1 | +5 | +10 |
|---|---|---|---|---|
| Transfer income (2022 NOK) | 2,621 | 12,947 | 22,968 | 24,566 |
| | (90) | (138) | (216) | (311) |
| Total income, all sources (2022 NOK) | -8,570 | -22,467 | -57,865 | -94,063 |
| | (1,356) | (1,558) | (4,319) | (2,829) |
| After-tax income (2022 NOK) | -4,491 | -11,336 | -30,371 | -51,173 |
| | (1,189) | (1,237) | (3,908) | (2,097) |
| Single | 0.0066 | 0.0431 | 0.0716 | 0.0858 |
| | (0.0005) | (0.0006) | (0.0009) | (0.0010) |
| Divorced or separated | 0.0035 | 0.0205 | 0.0367 | 0.0481 |
| | (0.0002) | (0.0003) | (0.0005) | (0.0006) |
| Childless | 0.0016 | 0.0030 | 0.0105 | 0.0161 |
| | (0.0003) | (0.0003) | (0.0005) | (0.0007) |

*Notes.* Event-time × treated coefficients, difference from event year −1. Standard errors clustered on the matched cell in parentheses. All six outcomes are estimated on the full sample.

## 6.1 Income Components

Transfer income rises by 2,621 NOK in the diagnosis year, 12,947 at +1, 22,968 at +5, and 24,566 at +10. Total income from all sources falls by 94,063 NOK at +10, and after-tax income by 51,173 (Figure 8).

Set against the gross earnings loss of 82,353 NOK at +10 (Table 2), the transfer increase offsets 29.8% of it. The transfer path rises steeply for five years and then flattens: the increment between +5 and +10 is 1,598 NOK, against 22,968 accumulated by +5. Transfers are therefore a smaller share of the earnings loss late in the window than early, because earnings continue to fall while transfers do not continue to rise.

This is direct evidence of welfare-state absorption, the channel highlighted in the introduction (cf. Bratsberg, Fevang and Røed 2010; Fevang, Hardoy and Røed 2017; Maestas, Mullen and Strand 2013). The Norwegian welfare state cushions but does not undo the earnings shock.

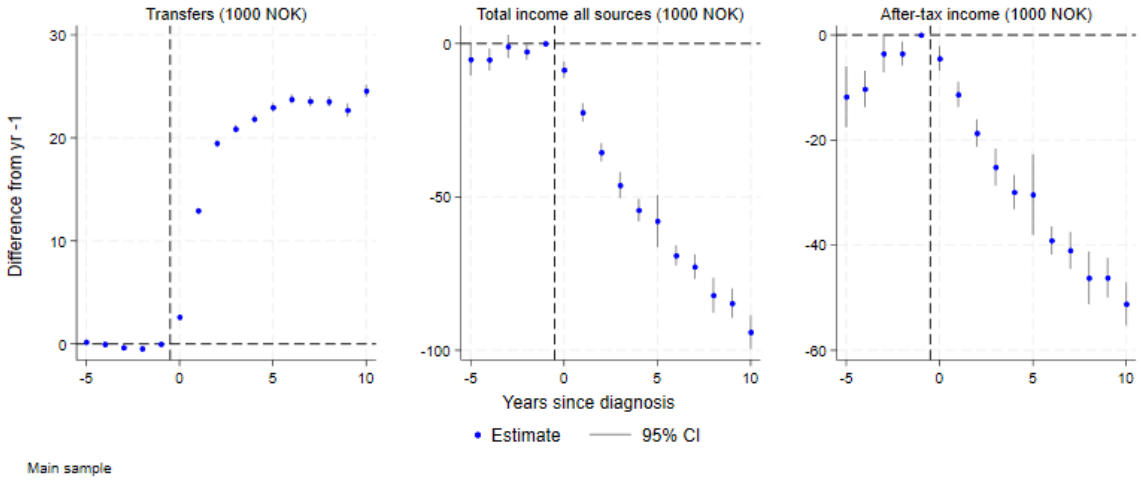



Figure 8: Event-study coefficients for transfer income, total income from all sources, and after-tax income (all in 1000 2022 NOK). Difference from year −1 with 95% CIs.

## 6.2 Family Outcomes

The treated are 8.6 percentage points more likely to be single at year +10 (0.7 points in the diagnosis year, 4.3 at +1), 4.8 points more likely to be divorced or separated, and 1.6 points more likely to be childless (Figure 9). Against a base-year childlessness rate of 0.32 among the treated, the last of these is a relative increase of about 5%.

The three differ in how clean the pre-period is. The divorce path is flat before diagnosis (leads of +0.0007, +0.0003, −0.0001, −0.0002 at years −5 to −2, none individually significant) and breaks at the diagnosis year. Being single has a small converging lead (+0.0062 at −5, falling to +0.0006 at −2), the same convergence toward the normalisation as the employment lead in Section 3.3, with the opposite sign. Childlessness sits between the two: its lead runs +0.0031 at −5 and crosses zero to −0.0018 by −2, so about a fifth of the eventual effect is present five years out. The path is not flat, though it is cleaner than a count of children, which we also examined.

The coupling and divorce results are large in relative terms and consistent with family-disruption mechanisms documented in the wider literature on health shocks and family stability (Fadlon and Nielsen, 2021; Hakulinen et al., 2020). The increase in childlessness is smaller in proportional terms.

These outcomes also bear on family spillovers: the diagnosed individual's partnership status is itself affected, which is upstream of any effect on the partner's labour supply or children's outcomes.

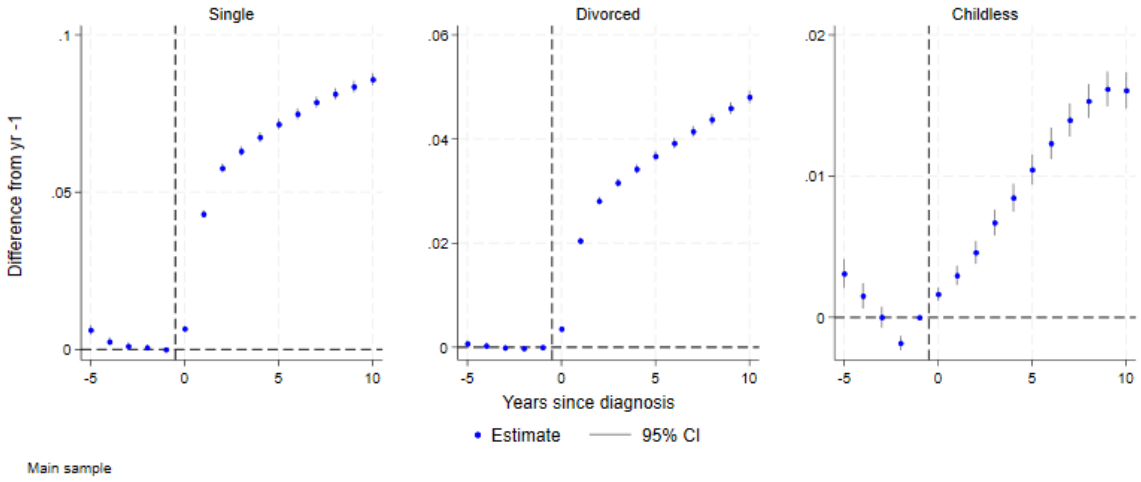



Figure 9: Event-study coefficients for being single, being divorced or separated, and being childless. Difference from year −1 with 95% CIs.

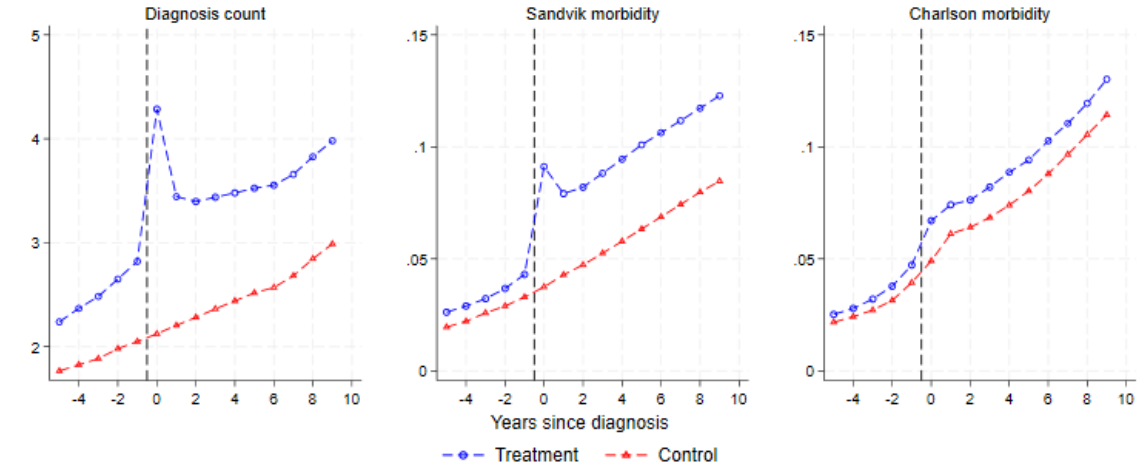
## 6.3 Post-Diagnosis Physical Morbidity

We now turn to physical morbidity as an outcome in its own right, measured after the mental health diagnosis rather than before it. We track three measures of somatic morbidity at each event time: the count of distinct non-MH ICPC-2 diagnoses over the past three years, the 24-category Sandvik morbidity index as a binary any-versus-none, and the Charlson comorbidity index computed from ICD-10 codes in the specialist register.

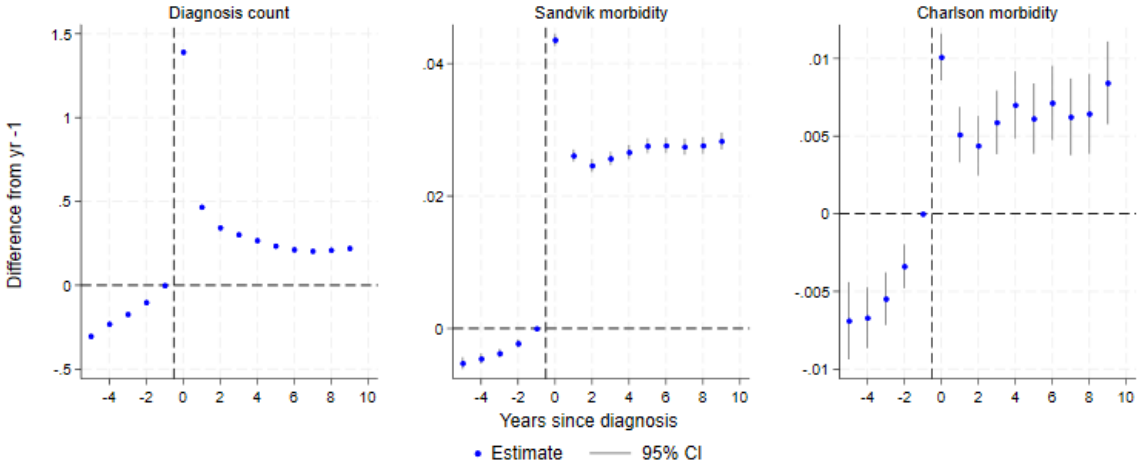
After MH diagnosis, the treated accumulate physical morbidity faster than matched controls on all three measures (Figure 10). At year +9, relative to year −1, the diagnosis count is 0.25 higher, following a contemporaneous spike of 1.4 at year 0 that partly reflects co-coding at the mental health visit itself. The Sandvik index is 0.03 higher and essentially flat after year 0, while the Charlson index is 0.008 higher and widens continuously across the event window.

The Sandvik and diagnosis-count series jump at year 0 (the MH-diagnosis visit picks up additional somatic conditions in the same encounter) and then stabilise at a positive offset. The Charlson series, by contrast, shows a smooth and widening gap, consistent with a genuine accumulation of severe somatic morbidity over time among the treated.

This result is direct evidence that an MH diagnosis is followed by genuine physical health deterioration, not just a relabelling of pre-existing complaints. It also speaks to one of the channels behind the persistent earnings decline.




(a) Means by event time, treatment vs control.




(b) Difference from year −1 (95% CIs).

Figure 10: Post-diagnosis somatic morbidity trajectories. Three measures: count of non-MH ICPC-2 diagnoses in past 3 years (left), Sandvik 24-category morbidity index (centre), Charlson comorbidity index (right). Treated accumulate physical morbidity faster than matched controls on all three measures.

## 6.4 Mortality

Mortality is estimated by Cox proportional-hazards on matched pairs (the survival-analysis commands do not accept weights, so the matched-pairs sample is used). On 886,002 pairs over a 14-year horizon, with 22,183 deaths, the hazard ratio for a first diagnosis is 1.258, with a 95% confidence interval of $[1.225, 1.292]$.

A first MH diagnosis is associated with a 26% higher mortality risk over the following 14 years. Kaplan–Meier survival curves, in Figure 11, diverge from approximately year 4 onwards.

This effect is large but plausible: it is consistent with the substantial excess mortality literature for severe mental illness (cf. Plana-Ripoll et al. 2023) and indicates that the persistent earnings decline coexists with a real health gradient, not just a labour-market disengagement.

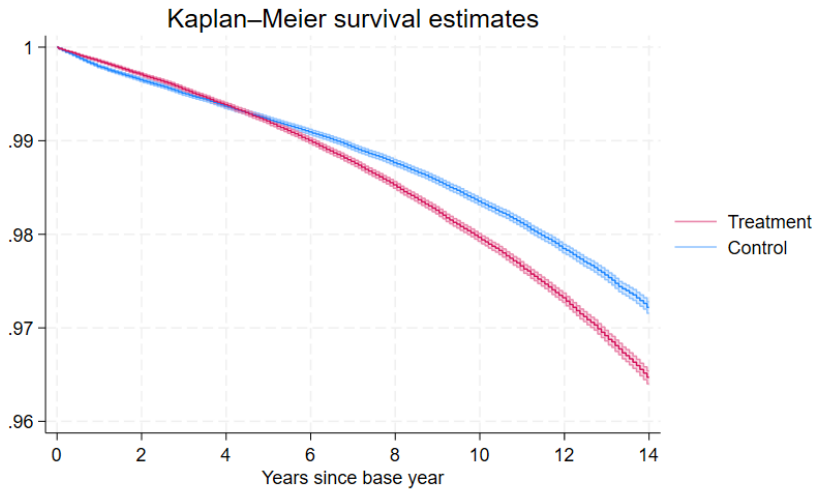



Figure 11: Kaplan–Meier survival, treatment vs. control, 14-year horizon (matched pairs, $N = 886{,}002$). Cox hazard ratio for treatment = 1.258 [1.225, 1.292]. Survival diverges from year ∼4 onwards.

## 6.5 Persistence of Mental Health Care Contact

The central interpretive question of the paper is why a first diagnosis is not followed by recovery in earnings. A natural first hypothesis is that the underlying illness simply persists. The registers let us test this directly: for every follow-up year we observe whether each individual has *any* MH diagnosis recorded in that year, and whether they carry the *same* diagnosis code as their initial one.

Table 7 reports the shares by event year. Three facts stand out.

Table 7: Mental health care contact after the first diagnosis.

| | Treated | | Comparison |
|---|---|---|---|
| Event year | Same diagnosis | Any MH diagnosis | Any MH diagnosis |
| +1 | 0.315 | 0.387 | 0.013 |
| +2 | 0.175 | 0.284 | 0.027 |
| +3 | 0.142 | 0.266 | 0.043 |
| +4 | 0.125 | 0.258 | 0.059 |
| +5 | 0.115 | 0.257 | 0.077 |
| +6 | 0.108 | 0.256 | 0.087 |
| +7 | 0.101 | 0.256 | 0.093 |
| +8 | 0.095 | 0.257 | 0.099 |
| +9 | 0.091 | 0.258 | 0.105 |
| +10 | 0.089 | 0.263 | 0.112 |

Shares of individuals with an MH diagnosis recorded in the given event year. "Same diagnosis" = the same code as the initial first diagnosis. Comparison individuals had no MH diagnosis as of the base year; their rising share reflects later first diagnoses. Cumulatively, 33.6% of the comparison group has received a first MH diagnosis by event year +10.

First, about a quarter of the treated remain in mental health care in any given year. The share with any MH diagnosis drops from 38.7% in the first year after diagnosis to 28.4% at +2 and about 26.6% at +3; and then stays essentially flat through year +10, ending at 26.3%. The share does not decline gradually; it stays at about a quarter, a stable stock of persistently ill individuals.

Second, persistence of the specific initial diagnosis is much lower: 31.5% carry the same code at +1, falling to 8.9% by +10. Roughly a third of the persistently-in-care group carries the same diagnosis year after year; the rest cycle through different MH diagnoses.

Third, the comparison group does not stay clean: by year +10, 11.2% of comparison individuals have an MH diagnosis in that year, and 33.6% have been diagnosed at some point. The within-year gap between treated and comparison narrows from 37 percentage points at +1 to 15 at +10. This is the contamination that motivates dropping future-treated person-years in the main specification (Section 3.1).

**Interpretation.** Reading the flat 26% as the persistently ill, roughly 60% of the treated eventually exit mental health care; they recover, clinically, at least as far as the registers can see. The persistent earnings decline therefore admits two non-exclusive readings: (i) the population-average effect is driven by the ∼25% who never recover, with magnitudes around four times the average for them and near zero for the recovered; or (ii) even those who exit care carry a lasting earnings scar. Distinguishing these is the single most important next step for the paper: a planned split of the event-study estimates by post-diagnosis persistence

status (and by employment status within the persistence table) will separate them.

The persistence facts also discipline the treatment-effectiveness discussion in Section 1: the stable 26% in-care share is consistent with meta-analytic estimates that a substantial minority of depression and anxiety cases are treatment-resistant or relapsing (Cipriani et al., 2018), and it gives the "why doesn't diagnosis trigger recovery?" question a concrete clinical anchor.

## 6.6 Occupational Rank

If mental illness erodes the quality of the match between worker and job, through reduced productivity, missed promotions, or moves into less demanding work, the earnings penalty should show up not only in employment exit but in the kind of jobs the treated hold. We test this by ranking roughly 450 four-digit occupations by their average earnings and assigning each employed person-year the rank of the occupation held.

Figure 12 shows mean occupational rank by event time, unconditional and conditional on employment. Both groups climb the occupational ladder over the 21-year window, as careers progress. But the treated climb more slowly after diagnosis: the treated–comparison gap is roughly 3 rank points through the pre-period (the treated sort into lower-paid occupations to begin with), is stable up to the base year, and then widens steadily to about 5 points by year +10. The pattern conditional on employment is just as strong; this is not a composition effect from low-rank workers exiting employment.

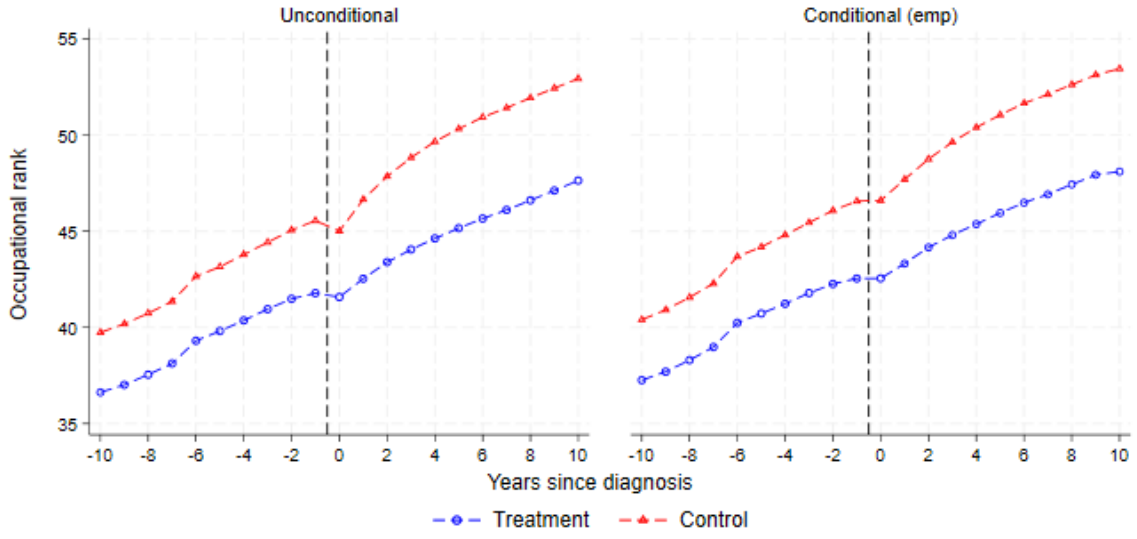



Figure 12: Mean occupational rank by event time, treated vs. comparison, unconditional (left) and conditional on employment (right). Occupations (∼450 at the 4-digit level) ranked by average earnings.

This is direct evidence for the reinforcing-feedback channel modelled by Jolivet and Postel-Vinay (2025) and Abramson, Boerma and Tsyvinski (2024): mental illness pushes workers down (or stalls them on) the job ladder, and worse jobs in turn are a worse platform for recovery. Together with the occupation-level nulls in Sections 5.3 and 5.4, no differential vulnerability by occupation characteristics, the picture is of a penalty that operates within every kind of job but accumulates through slowed career progression.

# 7 Conclusion

A first mental health diagnosis is followed by a large and lasting fall in employment and earnings. Ten years on, the diagnosed are 7.9 percentage points less likely to be employed and earn about 82,000 NOK less than their matched comparison individuals, with no sign of recovery at the end of the window. Transfers replace under a third of the lost earnings and stop growing after five years, while the earnings gap keeps widening.

We estimate the effect of a diagnosis, not of the underlying condition, and the two differ in both directions. The condition usually predates the first contact with the health system, so these estimates do not capture the full burden of the illness. At the same time, the diagnosis bundles the illness together with what the health system does about it: treatment, sick-listing, referral, and the record itself. What the design identifies is what happens to a career once the health system becomes aware of the condition, which is the margin policy can act on.

The absence of recovery does not contradict the evidence that treatment works. That evidence identifies narrow effects: a particular treatment, a particular diagnosis, or the selected group who reach specialist care, usually over a short horizon. We measure the population average across all diagnoses under routine care over a decade, and a treatable minority who improve is consistent with a flat average. Two of our own results point the same way. About a quarter of the treated are still in mental health care in every year through year +10. And those who stay employed fall about five rank points down the occupational ladder and do not climb back. Symptom relief and earnings recovery are not the same thing.

The uniformity of the decline is itself a finding. We see no meaningful differences by baseline somatic health, by occupational stress, by social-interaction intensity, or by sector and pay structure. Explanations that run through the demands of particular jobs, or through the interaction of mental with physical illness, predict gradients that are not there. What remains are mechanisms that operate on every kind of worker: the illness persisting, career scarring through slowed progression, and absorption into the transfer system. The exception is socioeconomic position, where losses are larger for the college-educated and for those whose fathers rank higher, the opposite of what a resource-based account predicts. Establishing whether that gradient is robust is the first task for further work.

# References


Abramson, Boaz, Job Boerma and Aleh Tsyvinski. 2024. "Macroeconomics of Mental Health." NBER Working Paper 32354.

Altonji, Joseph G., Daniel Giraldo Páez, Disa M. Hynsjo and Ivan Vidangos. 2021. Marriage Dynamics, Earnings Dynamics, and Lifetime Family Income. NBER Working Paper 28400 National Bureau of Economic Research. Revised November 2024.

Andersen, Signe Hald, Leah S. Richmond-Rakerd, Terrie E. Moffitt and Avshalom Caspi. 2024. "The Causal Effect of Mental Health on Labor Market Outcomes: The Case of Stress-Related Mental Disorders Following a Human-Made Disaster." *Proceedings of the National Academy of Sciences* 121(27):e2316423121.

Angelucci, Manuela and Daniel Bennett. 2024. "The Economic Impact of Depression Treatment in India: Evidence from Community-Based Provision of Pharmacotherapy." *American Economic Review* 114(1):169–198.

Angelucci, Manuela and Daniel Bennett. 2026. "Depression, Pharmacotherapy, and the Demand for a Preventive Health Product." Working paper, University of Texas at Austin and University of Southern California. Earlier version: IZA Discussion Paper 15832 (2022).

Angelucci, Manuela, Raissa Fábregas and Antonia Vazquez. 2026. The Well-Being Effects of Digital Mental Health Care. RFBerlin Discussion Paper 26107 ROCKWOOL Foundation Berlin. Also IZA Discussion Paper 18538.

Berk, Michael, Ole Köhler-Forsberg, Megan Turner, Brenda W. J. H. Penninx, Anna Wrobel, Joseph Firth, Amy Loughman, Nicola J. Reavley, John J. McGrath, Natalie C. Momen, Oleguer Plana-Ripoll, Adrienne O'Neil, Dan Siskind, Lana J. Williams, Andre F. Carvalho, Lianne Schmaal, Adam J. Walker, Olivia Dean, Ken Walder, Lesley Berk, Seetal Dodd, Alison R. Yung and Wolfgang Marx. 2023. "Comorbidity between major depressive disorder and physical diseases: a comprehensive review of epidemiology, mechanisms and management." *World Psychiatry* 22(3):366–387.

Bhalotra, Sonia, Adeline Delavande, Paulino Font-Gilabert and Joanna Maselko. 2025. "Maternal Investments in Children: The Role of Expected Effort and Returns." *The Economic Journal* 135(667):712–747.

Bhalotra, Sonia R., N. Meltem Daysal and Mircea Trandafir. 2025. Antidepressant Treatment in Childhood. CAGE Working Paper 766 Centre for Competitive Advantage in the Global Economy, University of Warwick. Also IZA Discussion Paper 18069 and CESifo Working Paper 12067.

Biasi, Barbara, Michael S. Dahl and Petra Moser. 2025. "Career Effects of Mental Health: Evidence from an Innovation in Treating Bipolar Disorder." *Journal of Political Economy Microeconomics* . Forthcoming.

Bjørnshagen, Vegar. 2021. "The Mark of Mental Health Problems: A Field Experiment on Hiring Discrimination Before and During COVID-19." *Social Science & Medicine* 283:114181.

Borusyak, Kirill, Xavier Jaravel and Jann Spiess. 2024. "Revisiting Event-Study Designs: Robust and Efficient Estimation." *Review of Economic Studies* 91(6):3253–3285.

Bratsberg, Bernt, Elisabeth Fevang and Knut Røed. 2010. "Disability in the Welfare State: An Unemployment Problem in Disguise?" IZA Discussion Paper 4897.

Callaway, Brantly and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225(2):200–230.

Cipriani, Andrea, Toshi A. Furukawa, Georgia Salanti, Anna Chaimani, Lauren Z. Atkinson, Yusuke Ogawa, Stefan Leucht, Henricus G. Ruhe, Erick H. Turner, Julian P. T. Higgins, Matthias Egger, Nozomi Takeshima, Yu Hayasaka, Hissei Imai, Kiyomi Shinohara, Aran Tajika, John P. A. Ioannidis and John R. Geddes. 2018. "Comparative Efficacy and Acceptability of 21 Antidepressant Drugs for the Acute Treatment of Adults with Major Depressive Disorder: A Systematic Review and Network Meta-Analysis." *Lancet* 391(10128):1357–1366.

Currie, Janet and Esmée Zwiers. 2025. "Medication of Postpartum Depression and Maternal Outcomes: Evidence from Geographic Variation in Dutch Prescribing." *Journal of Human Resources* 60(4):1093–1125.

de Quidt, Jonathan and Johannes Haushofer. 2016. Depression for Economists. NBER Working Paper 22973 National Bureau of Economic Research.

Deming, David J. 2017. "The Growing Importance of Social Skills in the Labor Market." *Quarterly Journal of Economics* 132(4):1593–1640.

Fadlon, Itzik and Torben Heien Nielsen. 2021. "Family Labor Supply Responses to Severe Health Shocks: Evidence from Danish Administrative Records." *American Economic Journal: Applied Economics* 13(3):1–30.

Fevang, Elisabeth, Inés Hardoy and Knut Røed. 2017. "Temporary Disability and Economic Incentives." *Economic Journal* 127(603):1410–1432.

Hakulinen, Christian, Marko Elovainio, Martti Arffman, Sonja Lumme, Kimmo Suokas, Sami Pirkola, Ilmo Keskimäki, Kristiina Manderbacka and Petri Böckerman. 2020. "Employment Status and Personal Income Before and After Onset of a Severe Mental Disorder: A Case-Control Study." *Psychiatric Services* 71(3):250–255.

Hakulinen, Christian, Marko Elovainio, Martti Arffman, Sonja Lumme, Sami Pirkola, Ilmo Keskimäki, Kristiina Manderbacka and Petri Böckerman. 2019. "Mental Disorders and Long-Term Labour Market Outcomes: Nationwide Cohort Study of 2,055,720 Individuals." *Acta Psychiatrica Scandinavica* 140(4):371–381.

Jolivet, Gregory and Fabien Postel-Vinay. 2025. "A Structural Analysis of Mental Health and Labour Market Trajectories." *Review of Economic Studies* 92(3):1920–1954.

Kessler, Ronald C., Steven Heeringa, Matthew D. Lakoma, Maria Petukhova, Agnes E. Rupp, Michael Schoenbaum, Philip S. Wang and Alan M. Zaslavsky. 2008. "Individual and Societal Effects of Mental Disorders on Earnings in the United States: Results from the National Comorbidity Survey Replication." *American Journal of Psychiatry* 165(6):703–711.

Laird, Jessica and Torben Heien Nielsen. 2016. "The Effects of Physician Prescribing Behaviors on Prescription Drug Use and Labor Supply: Evidence from Movers in Denmark." Job market paper, Harvard University.

Layard, Richard. 2013. "Mental Health: The New Frontier for Labour Economics." *IZA Journal of Labor Policy* 2:1–16.

Lund, Crick, Kate Orkin, Marc Witte, John H. Walker, Thandi Davies, Johannes Haushofer, Sarah Murray, Judy Bass, Laura Murray, Wietse Tol and Vikram H. Patel. 2024. "The Effects of Mental Health Interventions on Labor Market Outcomes in Low- and Middle-Income Countries." NBER Working Paper 32423.

Maestas, Nicole, Kathleen J. Mullen and Alexander Strand. 2013. "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt." *American Economic Review* 103(5):1797–1829.

Marcotte, David E. and Virginia Wilcox-Gök. 2003. "Estimating Earnings Losses Due to Mental Illness: A Quantile Regression Approach." *Journal of Mental Health Policy and Economics* 6(3):123–134.

OECD. 2012. Sick on the Job? Myths and Realities about Mental Health and Work. Technical report OECD Publishing Paris: .

Plana-Ripoll, Oleguer, Nanna Weye, Ann Kristin Knudsen, Christian Hakulinen, Kathrine Bang Madsen, Maria Klitgaard Christensen, Esben Agerbo, Thomas Munk Laursen, Merete Nordentoft, Allan Timmermann, Harvey Whiteford, Simon Øverland, Kim Moesgaard Iburg and John J. McGrath. 2023. "The Association Between Mental Disorders and Subsequent Years of Working Life: A Danish Population-Based Cohort Study." *Lancet Psychiatry* 10(1):30–39.

Ridley, Matthew. 2026. "Mental Illness Discrimination." Working paper, University of Warwick. Revise and resubmit, Journal of the European Economic Association.

Ridley, Matthew, Gautam Rao, Frank Schilbach and Vikram Patel. 2020. "Poverty, Depression, and Anxiety: Causal Evidence and Mechanisms." *Science* 370(6522):eaay0214.

Shapiro, Bradley T. 2022. "Promoting Wellness or Waste? Evidence from Antidepressant Advertising." *American Economic Journal: Microeconomics* 14(2):439–477.

Sun, Liyang and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225(2):175–199.

# A Additional Tables and Figures

Table A1: Register coverage of the population aged 20–65

| Year | Pop 20–65 | Emp 20–65 | KUHR ever | NPR ever | NPR first / KUHR ever |
|---|---|---|---|---|---|
| 2008 | 2,426k | 1,952k | 552k | 98k | 84k |
| 2013 | 2,405k | 1,917k | 938k | 314k | 30k |
| 2018 | 2,379k | 1,883k | 1,191k | 480k | 25k |
| 2023 | 2,329k | 1,876k | 1,394k | 620k | 22k |

Table A2: Callaway–Sant'Anna dynamic ATT estimates, real earnings (2022 NOK), 2012 base-year test cohort.

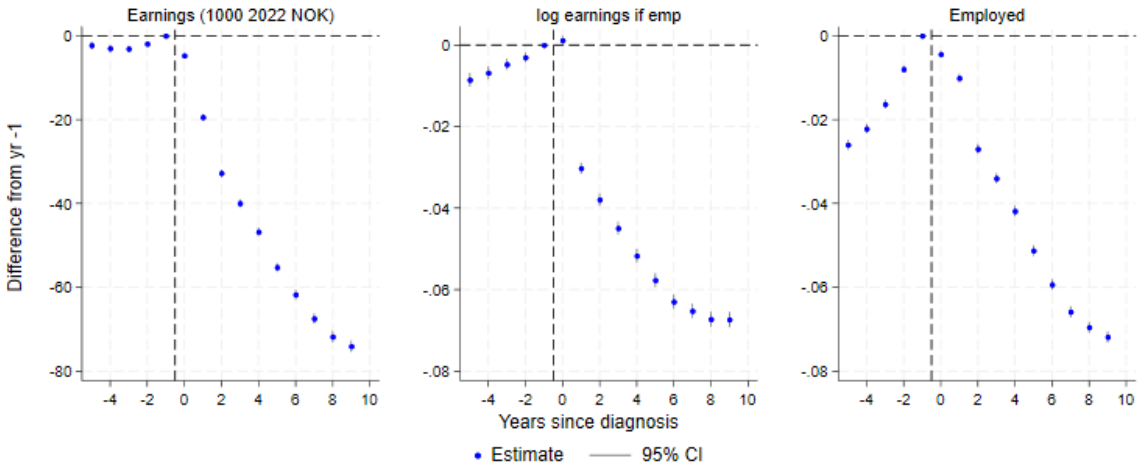
| Event year | ATT | Std. err. |
|---|---|---|
| 0 | −9,667 | 602 |
| +1 | −33,504 | 820 |
| +2 | −50,358 | 1,071 |
| +3 | −57,140 | 1,209 |
| +4 | −62,980 | 1,386 |
| +5 | −70,239 | 1,695 |
| +6 | −73,972 | 1,795 |
| +7 | −78,300 | 2,018 |
| +8 | −83,772 | 2,215 |
| +9 | −87,819 | 3,205 |
| +10 | −91,520 | 2,698 |
| Post average | −63,570 | 1,347 |

Control group: not-yet-treated. Outcome model: regression adjustment. $N = 3{,}642{,}514$ person-years.

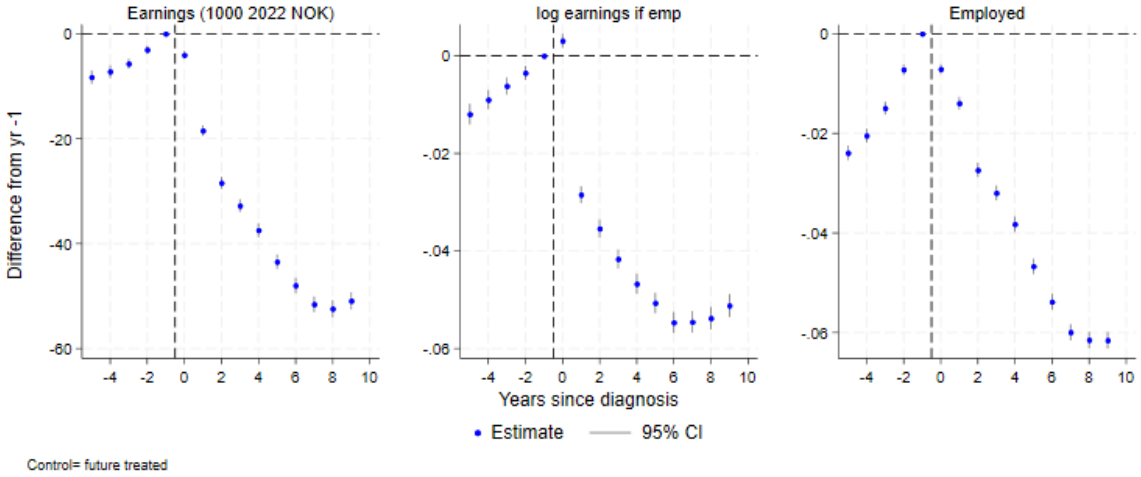
Table A3: Year-+10 earnings effect (2022 NOK) across specifications, full June panel.

| Specification | Earnings at +10 |
|---|---|
| Main (individual FE + base-year × event-time FE) | −85,916 |
| restricted to employed 5 years pre-base | −73,988 |
| restricted to positive wage 5 years pre-base | −73,075 |
| balanced panel (observed all 21 event years) | −69,178 |
| Matched pairs, demographic | −98,677 |
| Matched pairs, demographic + industry + occupation | −86,171 |

All specifications weight by the demographic balance weights and cluster on the individual (matched-pairs designs cluster on the pair).




(a) Reweighted population (age, sex, base year).




(b) Future-treated as control group.

Figure A1: Alternative identification strategies, 5-pre/9-post window. Difference from year −1 with 95% CIs.

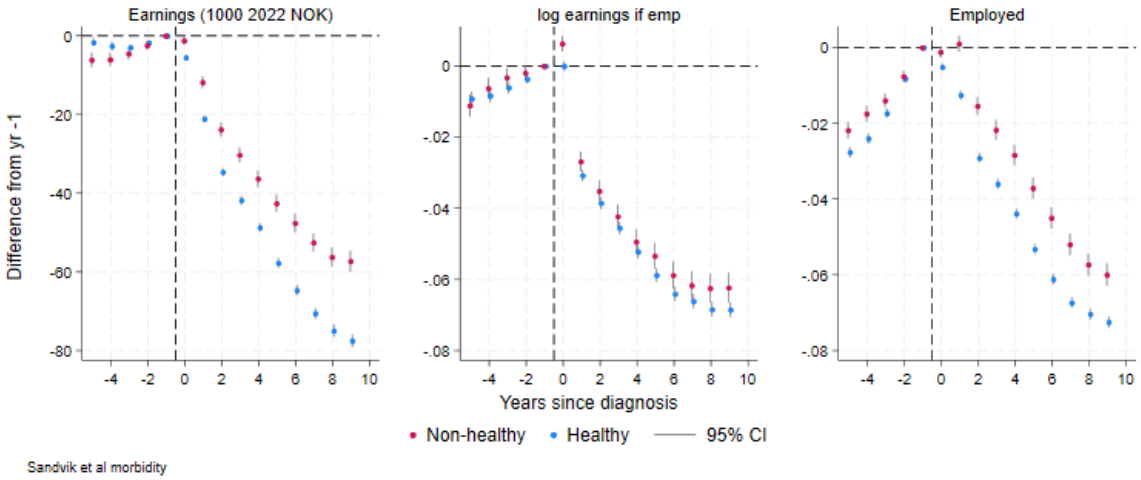



Figure A2: Event-study coefficients by baseline somatic morbidity (Sandvik-based 24-category binary index, 3-year lookback in KUHR; P-chapter codes excluded). Non-healthy = at least one somatic morbidity category (20% of the sample). In absolute earnings, healthier individuals show slightly larger drops; in log earnings and employment, effects are essentially identical across groups.

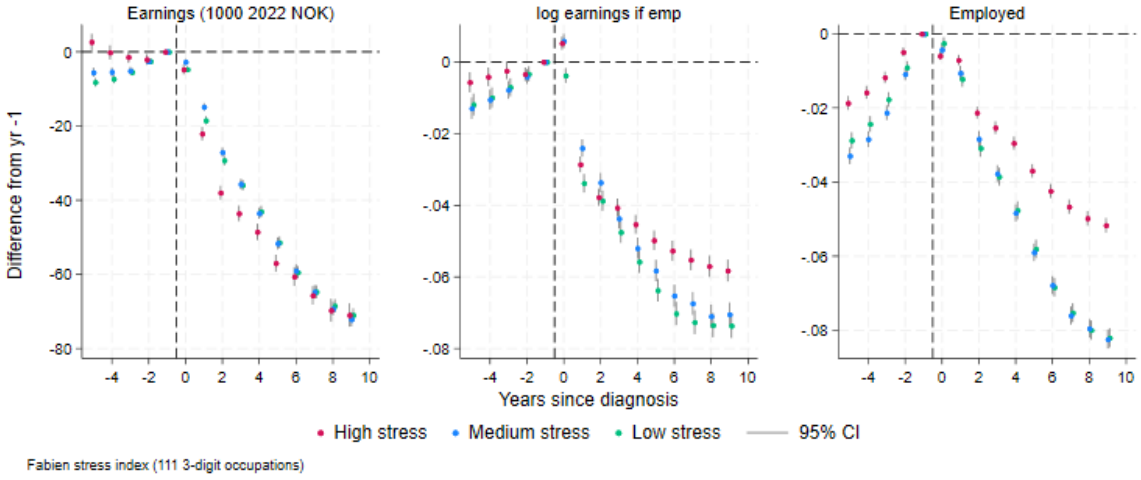



Figure A3: Event-study coefficients by occupational stress tertile, broad index (principal component of 29 O*NET items, 111 ISCO-88 three-digit occupations). All-sample specification. Tertiles are essentially indistinguishable in log earnings and employment; the high-stress tertile shows a slightly smaller absolute earnings drop at year +9.

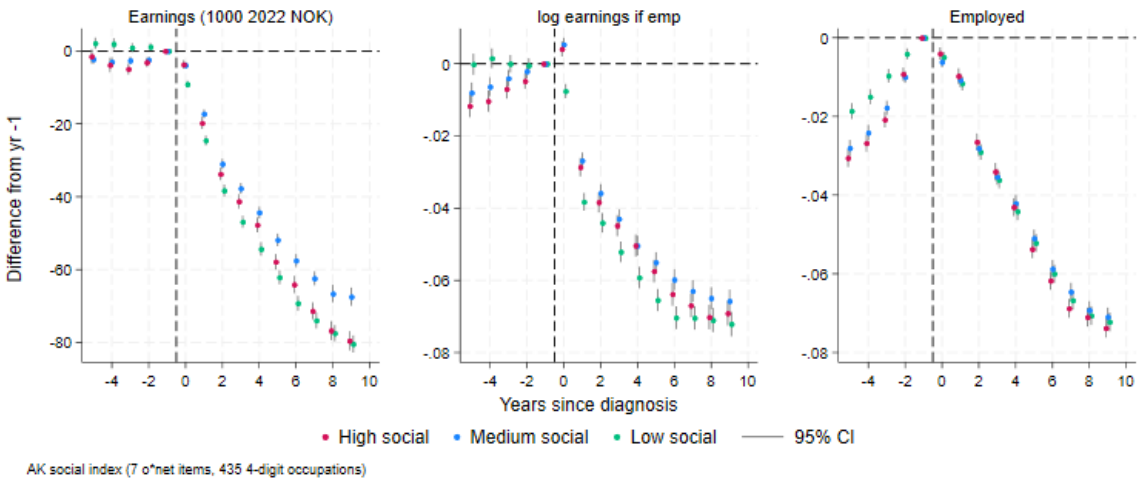



Figure A4: Event-study coefficients by social-interaction tertile, narrow index (seven O*NET items, 435 ISCO-88 four-digit occupations). Within-firm specification (workplace fixed effects, firm-balance weights). Tertiles are indistinguishable in all three outcomes.

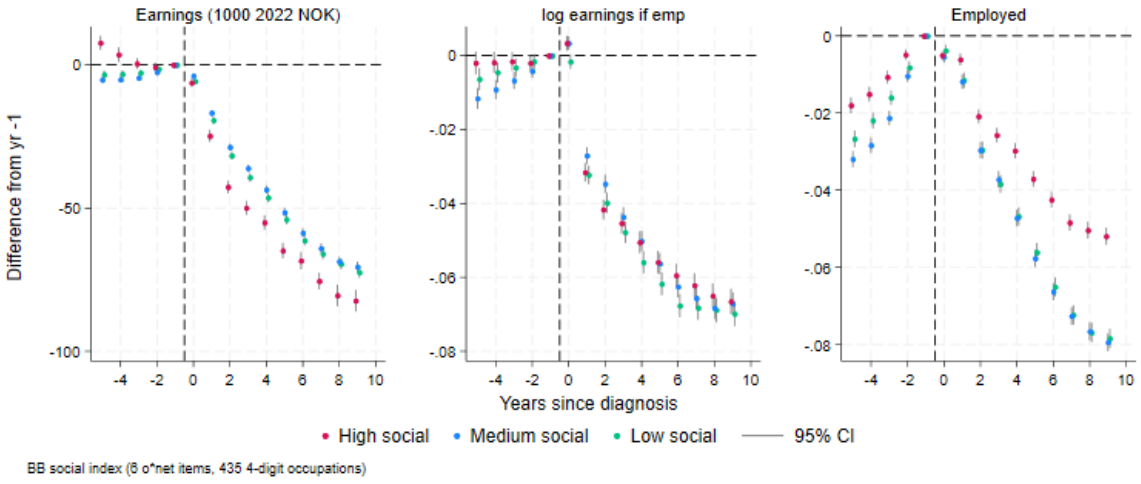



Figure A5: Event-study coefficients by social-interaction tertile, broad index (six O*NET items). Within-firm specification. Same flat pattern as the narrow index.

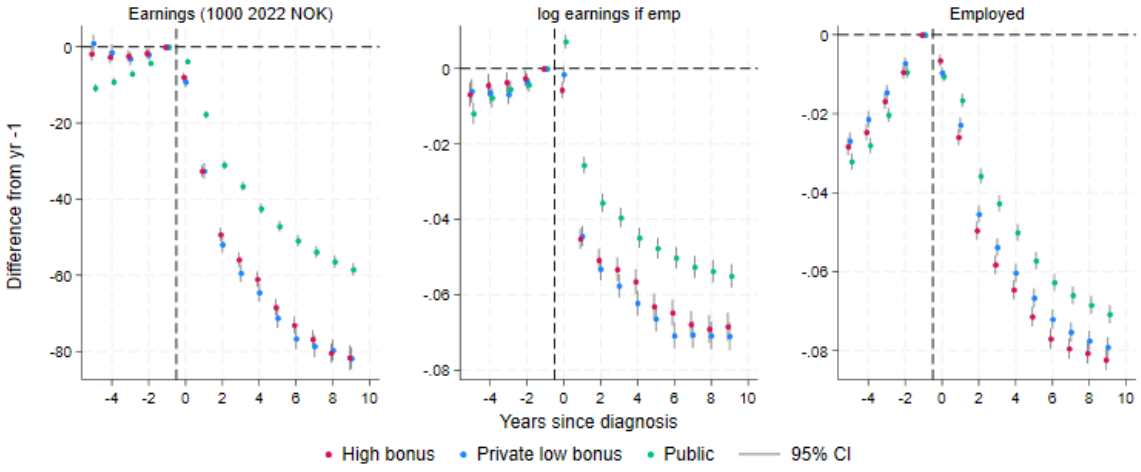



Figure A6: Event-study coefficients by sector and within-private bonus intensity; earlier-round sample (base years 2009–2013). Public sector earns a smaller earnings and log-earnings drop than either private-sector category; the two private series are indistinguishable.

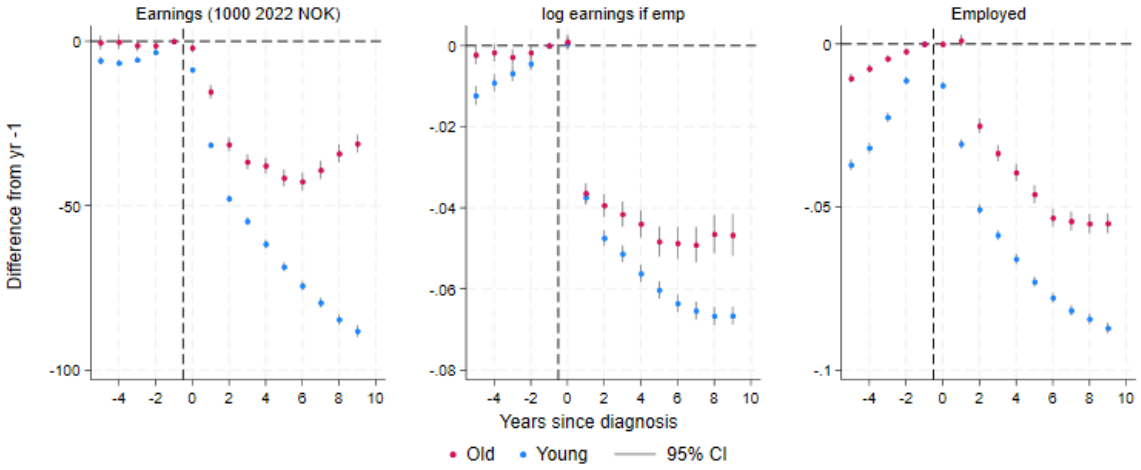



Figure A7: Event-study coefficients for MH by age at diagnosis. Difference from year $-1$ with 95% CIs. Young workers ($< 50$ at diagnosis) lose roughly three times more in absolute earnings than older workers ($\geq 50$).