(Not) Talking Through Mental Health Problems

How Talk Therapy Impacts Healthcare Utilization, Health and Employment

Roger Prudon*

March 22, 2026

Abstract

To increase the uptake of mental health care and reduce the burden of mental health, many countries have introduced programs offering easily accessible talk therapy. Despite their widespread introduction and continued expansion, causal evidence on these programs' effectiveness is lacking. I estimate the impact of a Dutch program that introduced talk therapy provided by nurse practitioners at GP practices nationwide. The reform increased talk therapy uptake by 6 percentage points among individuals seeking care but also widened existing disparities in care utilization. The increased take-up of talk therapy did not significantly improve the health and employment outcomes of patients.

Keywords: Mental Health Care, Talk Therapy, Employment, Health Care Inequalities

JEL: J08, I1, J22

*Affiliation: Lancaster University. Contact: r.prudon@lancaster.ac.uk. I am grateful for useful comments from Amanda Kreider, Christoph Kronenberg, Coleman Drake, Shushan Margaryan, and Shooshan Danagoulian and from participants at the Essen Health Economics Seminar Series, LoLa HESG, American-European HESG, KVS New Paper Sessions, RES, IHEA, EEA-ESEM, Rodolfo Benedetti conference, and EMH workshop. I gratefully acknowledge financial support from NWO for this research project. The results in this paper are based on nonpublic microdata from Statistics Netherlands (CBS) and Nivel. In principle, these microdata are accessible for statistical and scientific research (further information: microdata@cbs.nl). This study has been approved according to the governance code of Nivel Primary Care Database, under number NZR-00324.04. The use of electronic health records for research purposes is allowed under certain conditions. When these conditions are fulfilled, neither obtaining informed consent from patients nor approval by a medical ethics committee is obligatory for this type of observational study containing no directly identifiable data (art. 24 GDPR Implementation Act jo art. 9.2 sub j GDPR).

1 Introduction

According to recent estimates, approximately one in eight people worldwide live with a mental health disorder (WHO, 2022), which imposes a significant health burden on society. Given the impact that mental health disorders can have on individuals' labor market outcomes, this mental health burden implies substantial economic losses. In 2010, the economic burden on the world economy from poor mental health was estimated at $2.5 trillion (approximately 5% of world GDP)—a figure predicted to grow to $6 trillion by 2030 (Lancet, 2020).

The medical literature has established the effectiveness of treatments in reducing the health burden of mental health disorders (Hofmann et al., 2012). Treatment could thus potentially lead to significant gains for society. However, a large share of the affected individuals do not receive (adequate) treatment, with estimates indicating that 30% to 70% of people with a mental health disorder remain untreated (Kohn et al., 2004; Alonso et al., 2018). This is especially true among disadvantaged groups such as those with a migration background (Sentell et al., 2007).

In the last two decades, in response to this low treatment uptake, many countries have introduced programs that provide easily accessible forms of mental health care. Examples include Improving Access to Psychological Therapies (now called NHS Talking Therapies; United Kingdom), the Better Access Initiative (Australia), the Ontario Structured Psychotherapy Program (Canada), Prompt Mental Health Care (Norway), and Mental Health in Primary Care (Netherlands). In the US, similar care is often provided through community mental health centers or private organizations such as Sondermind or the Open Path Collective.

Although these programs differ in their exact setup, they all provide a relatively low-intensity form of mental health care at low cost with as few barriers to take-up as possible.[1] By facilitating treatment access, they aim to reduce disparities in mental health care by increasing uptake among underserved groups. The type of treatment provided through these programs is often referred to as *talk therapy*, an umbrella term for various

[1] Appendix Table A.1 compares the main characteristics of the various programs.

forms of psychotherapy such as cognitive behavioral therapy and cognitive analytical therapy. Since their inception, these programs have grown rapidly: Every year, over 500,000 patients in each of the UK and the Netherlands and 1.3 million patients in Australia (1–3% of the population) receive care under such initiatives.

Despite their proliferation, evidence of these programs' effectiveness is limited. This paper aims to fill this gap by examining the impact of the introduction of a program offering easily accessible talk therapy provided by nurse practitioners specialized in mental health care (NPMHs) at general practitioner (GP) practices in the Netherlands. I first determine to what extent the program's introduction impacted take-up of the new treatment itself and of all other types of mental health care. Next, I examine whether the increase in mental health care utilization improved the health and labor market outcomes of individuals suffering from mental health disorders.

To estimate the causal impact of the program, I exploit exogenous variation in the probability of receiving talk therapy, driven by the introduction and gradual increase in funding for this type of talk therapy in the Netherlands. The primary analysis employs an instrumental variable (IV) strategy that exploits variation in GPs' propensity to refer patients to this new treatment. This method thus effectively compares similar patients facing different probabilities of receiving talk therapy as GP practices rolled out the new treatment. As a robustness check, I use a regression discontinuity (RD) design to compare individuals seeking mental health care right before and right after their GP practice first hired a nurse practitioner to provide talk therapy.

Using both methods, I find that the introduction of talk therapy at GP offices significantly increased take-up of this treatment. Patients contacting their GP for mental health–related reasons were approximately 6 percentage points (pp) likelier to receive talk therapy after the introduction of this new treatment. However, the impact on receipt of basic/specialized mental health care and/or pharmaceutical treatment in the short and long run (up to 10 years) was very limited, pointing to limited preventive health care gains of talk therapy.

Given the substantial increase in uptake of talk therapy, I next examine whether the

introduction of talk therapy in GP practices reduced existing disparities in health care utilization. To this end, I first estimate health care disparities using administrative data on health care utilization as Sentell et al. (2007) and Cook et al. (2017) do for the US. Whereas previous literature has estimated these disparities by examining only utilization rates, I combine measures of utilization and need for care. Controlling for the need for care substantially impacts both the magnitude and the sign of the estimated disparities. Combining data on health care need and utilization, I find that older individuals, men, and people with a migration background or lower educational attainment used mental health care disproportionately less before the introduction of talk therapy at GP practices.

With these disparities in mind, I estimate heterogeneity in the impact of the talk therapy rollout by background characteristics of the individuals seeking care. Individuals under the age of 45, women, and people without a migration background or with higher educational attainment were likelier to take up this new treatment. These groups already had relatively high utilization of mental health care before the program was implemented, which implies that the increase in the use of talk therapy worsened existing health care disparities.

Finally, I examine whether the increase in uptake of talk therapy improved the health and employment outcomes of individuals seeking care. Despite the substantial increase in uptake, I find no significant impacts on a wide range of patient outcomes. The increased uptake of talk therapy did not lead to significant improvements in self-reported or physician-assessed measures of (mental) health in the long run. Additionally, patients were no likelier to be employed or to have higher labor earnings, nor were they any less likely to receive social benefits such as sickness or disability benefits. This was the case in both the short and the long run, with labor market outcomes remaining unaffected for up to 10 years after treatment commenced.

This paper is most closely related to a strand of literature examining the introduction of similar programs in the Netherlands and other countries. Dros et al. (2024) show that in the Netherlands, patients at the GP practices likelier to refer patients to NPMHs have lower utilization rates of basic and specialized mental health care. Similarly, Clark et

al. (2018) find significant correlations between characteristics of the UK program (waiting times, number of sessions, etc.) and mental health outcomes. Oparina et al. (2024) and Rzepnicka et al. (2025) use patients who are waitlisted or drop out of treatment in the UK program as a control group for those receiving treatment and find that the program significantly increases recovery rates and improves employment outcomes. Using a prospective cohort design, Sæther et al. (2019) report long-lasting and significant improvements in mental health outcomes among patients of the Norwegian version of this program. In contrast to these studies finding positive impacts, correlational studies examining population-level utilization rates and mental health outcomes in the UK and Australia find no significant impact of these programs (Jorm, 2018; Bastiampillai et al., 2019).

The estimates from the studies mentioned above should be interpreted with caution, as correlational studies such as Dros et al. (2024), Jorm (2018), Bastiampillai et al. (2019), and Clark et al. (2018) cannot capture the dynamic interplay between *need* for mental health care and the potential policy responses that accompany it. Additionally, designs using waitlisted patients as controls, as in Oparina et al. (2024), have received considerable criticism in recent years given the potential direct impact of waiting times on the control group (Bruun et al., 2022). Similarly, designs that include patients who have dropped out of treatment, as in Rzepnicka et al. (2025), can yield biased estimates because the decision to drop out may be driven by factors such as the severity of the mental health problems. I confirm these concerns by applying the methodology of Rzepnicka et al. (2025) to the Dutch setting and showing that it yields biased estimates.

Therefore, to my knowledge, this paper is the first to establish the causal impacts of the introduction of an easy-access form of talk therapy on patient outcomes. While I find significant increases in talk therapy uptake, the absence of impacts on other types of mental healthcare and measures of (mental) health itself indicates that the mental health gains are limited, in contrast to what much of the prior literature has found.

This paper additionally extends the literature on similar programs by expanding the scope to outcomes beyond mental health (care) and incorporating patient labor market

outcomes. The findings therefore also contribute to the literature on the causal impact of mental health treatment on employment outcomes. This literature has predominantly found significant improvements in labor market outcomes from mental health treatment. Biasi et al. (2021) find that lithium reduces the negative earnings penalty of bipolar disease by one-third, or approximately 10 pp. Similarly, Prudon (2025) finds that among individuals who have to wait one additional month before treatment commences, the employment rate drops by 2 pp. Last, Shapiro (2022) finds that a 10% increase in advertising for antidepressants leads to a 0.3% rise in prescriptions and a reduction in the number of missed days at work. I find, in contrast, that talk therapy provided by nurse practitioners does not significantly impact employment, earnings, or sickness benefit uptake.

Finally, this paper relates to the strand of literature examining differences in mental health treatment uptake. Studies on the US have identified significant disparities in treatment uptake, with disadvantaged groups being up to 80% less likely than more advantaged patients to seek mental health care. Language barriers, such as limited English proficiency, are a primary factor in this reduced access (Sentell et al., 2007; Cook et al., 2017). My paper makes three crucial contributions to this literature. First, I estimate differences in uptake of mental health care in the Netherlands and find, in contrast to Sentell et al. (2007) and Cook et al. (2017), that disadvantaged groups such as those with a migration background are likelier to use care than individuals without a migration background. However, as a second contribution, I show that accounting for the need for mental health care can significantly affect both the magnitude and the sign of estimated disparities. Third, I show that the introduction of an easy-access treatment—in this case, talk therapy offered at GP practices—does not necessarily decrease disparities in health care utilization. While the new talk therapy program increased treatment uptake among all subgroups, the groups already likeliest to use mental health care *before* the program's introduction experienced the largest increase. The roll-out of this program thus widened existing treatment gaps based on the age, gender, education level, and migration background of the individuals suffering from mental health problems.

In short, the introduction of talk therapy at GP practices in the Netherlands significantly increased treatment uptake. This increase was concentrated among groups that already had a relatively high probability of using mental health care, thereby potentially further increasing existing disparities. While the increased take-up of treatment may have improved patient well-being, it did not reduce future mental health care need or improve patient health and labor market outcomes up to 10 years after the start of treatment. The results thus highlight that, given the breadth of mental health disorders and the variability in treatment options available, the effectiveness of treatment in terms of future health (care) gains and labor market improvements is not guaranteed. Additional efforts are required to reduce the (economic) impact of mental health problems and close disparities in mental health care utilization.

This paper proceeds as follows. The Dutch institutional setup of mental health care and the talk therapy reform are discussed in Section 2. Section 3 presents the data used and shows descriptive statistics for the sample of interest. The main empirical results on the impact of talk therapy are presented in Section 4. Finally, Section 5 concludes.

2 Mental health care in the Netherlands

2.1 Prior to the reform

In the Netherlands, GPs are the gatekeepers of the health care system. If an individual has health issues, she must contact her GP to receive a referral for further care. In the case of mental health problems, GPs make a first assessment of the severity of the condition; they then have various treatment options they may refer the patient to. If a GP deems the mental health problem very limited, he can decide not to refer the patient for any treatment and may ask the patient to return after a specific period to check on the mental health problem. If the problem is more severe, the GP can refer the patient for either basic or specialized mental health care. Basic mental health care is focused on "singular" mental health disorders, whereas specialized mental health care is aimed at more complex cases. Alongside these various referral options, GPs can decide

to prescribe pharmaceuticals. Medical guidelines, however, state that they should do so "with caution."

All costs of care provided by GP practices are covered by the mandatory basic health insurance system, with no copayments or deductibles. If GPs refer patients for specialized care, copayments might be required. Additionally, there is an annual deductible of €385 for specialized care.[2] Given the extensive coverage of the basic health insurance system and its mandatory nature, private provision of care in the Netherlands is limited. Except in urgent situations, individuals can attend only the GP practice where they are registered, and the practice must be located close to where they live.

2.2 After the reform

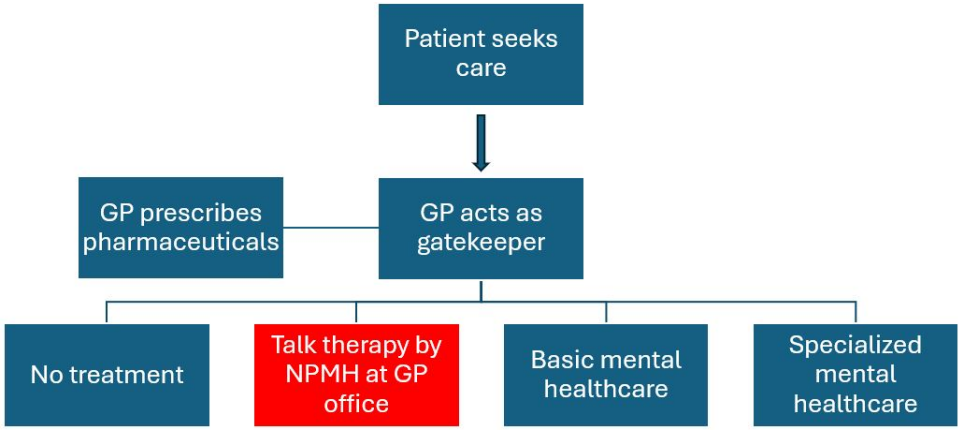
To increase mental health treatment uptake and alleviate pressure on the basic and specialized mental health care systems, the Dutch government decided to expand the treatment options available to patients. Since 2008, GP practices have been able to hire NPMHs. These nurse practitioners are most often not medical doctors but are required to have a 4-year nursing degree (or equivalent) and at least 2 years of mental health-related work experience. Individuals who meet these requirements can complete a one-year program to receive their NPMH accreditation.

As NPMHs are not medical doctors, they are not allowed to prescribe pharmaceuticals. They do, however, provide various forms of talk therapy. The treatment provided by NPMHs is offered at the GP practice, which implies that, in contrast to basic and specialized mental health care, patients need not make any copayments when being treated by an NPMH. Treatment by NPMHs is intended mainly for patients with mild mental health disorders who, prior to the program's introduction, would have been on the margin between receiving no treatment and being referred for basic mental health care. Figure 1 illustrates the various treatment options available since the introduction of NPMHs, ordered (from left to right) by the severity of the health problems to be treated.

Funding for talk therapy at GP practices was first made available in the Netherlands

[2]The annual deductible has increased over time, and individuals can opt for a higher deductible of up to €885 in exchange for a reduction in their health insurance premium.

Figure 1: Treatment options available after the introduction of NPMHs

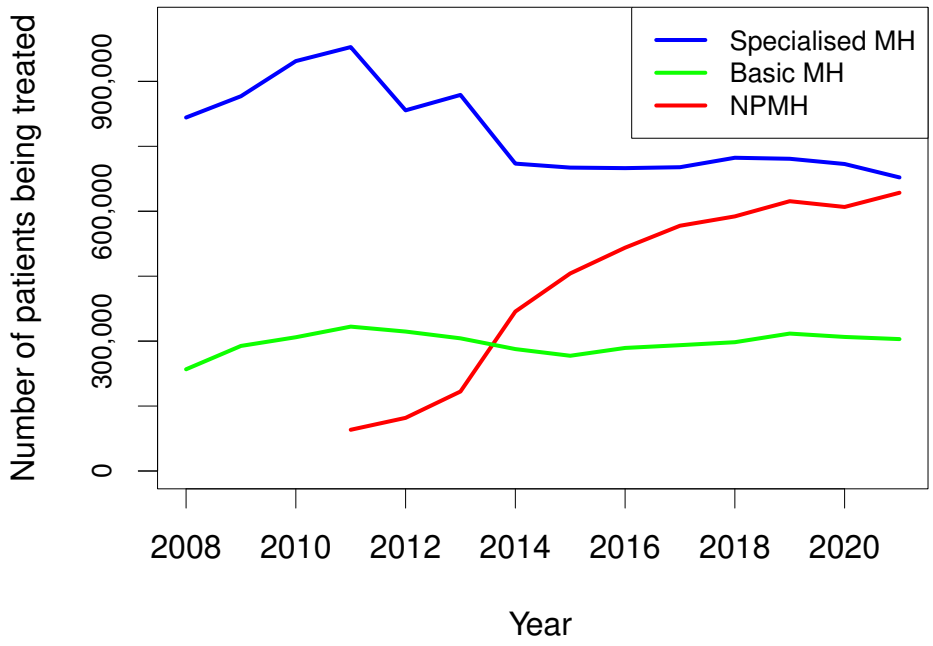


in 2008. Initially, GP practices could receive funding for 4.5 hours (0.125 FTE) of NPMH labor per week for every 2,350 patients registered at the practice. Since 2008, this funding steadily increased to 16 hours (0.5 FTE) of NPMH labor per week for every 2,350 patients in 2024. As the amount of funding available has risen, the number of GP practices that offer this type of treatment has also increased. To illustrate the gradual expansion of treatment by NPMHs, Figure 2 shows the number of patients being treated each year in basic or specialized mental health care and by NPMHs at GP practices. The numbers of patients in both basic and specialized mental health care were relatively stable over the focal period at 300,000 and 700,000, respectively. Meanwhile, the number of patients being treated by NPMHs grew substantially from 95,000 in 2011 to over 600,000 in 2021.[3]

The introduction of talk therapy at GP practices had several aims. The primary aim was to increase treatment uptake, especially among disadvantaged groups with relatively low mental health care utilization rates. Second, it was meant to alleviate pressure on other treatment providers by treating patients who would otherwise require treatment from their GP or basic/specialized mental health care providers. Last, it aimed to improve patients' long-term outcomes by means of increasing treatment uptake. This was expected to be achieved at relatively low cost, as treatment by NPMHs is cheaper than basic/specialized mental health care. The conclusion (Section 5) returns to these aims and discusses the extent to which they have been achieved.

[3]Note: Information on NPMH utilization for 2008-2010 is not available.

Figure 2: Number of patients being treated per year for each type of mental health care treatment



3 Data

To assess the impact of talk therapy provided at GP practices, I merge several administrative data sources from various providers to obtain individual-level information on health care utilization, labor market outcomes, and background characteristics. All datasets are linked via anonymized citizen service numbers. I first discuss the various datasets and then show descriptive statistics in Section 3.4.

3.1 Data on health care provided at GP practices

Detailed data on health care provided at GP practices are, unfortunately, not available in the standard Dutch administrative datasets. I therefore use data from Nivel, an association of GP practices in the Netherlands. Approximately 600 GP practices, covering 10% of the entire Dutch population, were affiliated with Nivel in the study period. While the sample is not random, the population served by Nivel-affiliated GP practices is representative of the Dutch population on a wide range of background characteristics.

Nivel has provided data on all mental health–related GP consultations between 2011

and 2021. For these consultations, I observe patient and GP practice identifiers, the exact date, and the reason for the consultation based on International Classification of Primary Care (ICPC) codes. In addition, Nivel has provided the same data for all NPMH consultations. These data make it possible to construct time series for treatment by NPMHs at both the individual and the practice levels.

Note that GP practices could (and in some cases did) start their affiliation with Nivel after the start of the observation window (2011) or terminate their affiliation during this window. Additionally, some patients changed their GP upon relocating to a different municipality. For both of these types of cases, data on health care provided at the GP practice are not available.

3.2 Data on (mental) health care

The data on health care provided at GP practices are complemented with a dataset provided by Statistics Netherlands showing annual health care expenditures covered by the basic health insurance system from 2009 to 2020. These expenditures are categorized into subcategories, which enables me to distinguish expenditures at GP practices (which combine both GP and NPMH care), expenditures for basic mental health care and specialized mental health care, spending on pharmaceuticals, and spending on other types of health care (non–mental health care).[4] I use this data source to create an annual time series of utilization of the various types of health care.

While covering all types of health care, the data on expenditures are available only at annual level and have limited detail. I therefore use an additional dataset obtained through the health insurance system, which contains more granular information on specialized mental health care use. Specifically, this data source contains information on all treatment-related events for specialized mental health care that took place between 2011 and 2021. Mental health care treatment is defined as any in-person or virtual interaction with a mental health care provider. For each event, the dataset includes details such as the exact event date, event type (e.g., initial contact, intake, treatment, administrative),

[4]See Appendix Table A.3 for details on the health care spending categories.

duration of patient contact in minutes, mental health diagnosis, category of treatment provider (psychologist, psychiatrist, psychotherapist, other), and anonymized identifiers for both the patient and the provider. For a subset of all treatment events, the data also contains a measure of the severity of symptoms as assessed by the physician.

The health care data are supplemented with detailed information on the use of pharmaceuticals for mental health–related diagnoses. This dataset includes all individual prescriptions covered by the basic health insurance system for a selection of ATC-4 code medications, along with the dates the prescriptions were filled. The data cover seven drug categories: (1) antidepressants, (2) antipsychotics, (3) anxiolytics, (4) hypnotics, (5) psychostimulants, (6) medications for addictive disorders, and (7) opioids.[5] Prescriptions are included in the dataset if issued by a GP or mental health care provider. Combining these four datasets yields individual-level monthly time series on all available treatment options for mental health care in the Netherlands for a sample of individuals who contacted their GP for a mental health–related reason.[6]

Lastly, to measure disparities in health care utilization, I require measures of both health care utilization and health care need. To this end, I merge a large survey with the administrative datasets. This survey, the National Health Monitor (Gezondheidsmonitor), has been administered by Statistics Netherlands every four years since 2012 and covers a representative sample of approximately 400,000 respondents (2% of the Dutch population). The survey contains a wide range of health and health care questions. Most importantly for this paper, it includes questions about self-reported mental well-being and measures that summarize the risk of the most frequent mental health disorders, such as anxiety and depression.[7] Combining the information on health care utilization and the survey measures on health care need yields a proxy of the gap in treatment uptake for various groups of the population. I use this proxy in the heterogeneity analysis to determine whether the introduction of talk therapy at GP practices narrowed or widened health care disparities.

[5]The corresponding ATC-4 codes are N02A, N05A, N05B, N05C, N06A, N06B, and N07B.

[6]The only exception to this is basic mental health care, for which only annual time series are available.

[7]The survey contains both the responses to the underlying questions used to compute these risk measures and the risk measures themselves.

3.3 Data on labor market outcomes and other characteristics

Finally, the health care data are merged with data on labor market outcomes. These include monthly metrics for employment, labor earnings, working hours, and receipt of unemployment benefits, sickness/disability benefits, and social assistance. The labor market panel covers 2006–2023, with data on monthly labor earnings and working hours available from the years from 2009. For most analyses, the time series are aligned relative to patients' initial contact with a GP. Individuals are tracked from eight years before this initial contact to ten years after; this results in an unbalanced panel.

The panel data on health and labor market outcomes are supplemented with administrative records from Statistics Netherlands, which include information on date of birth, gender, migration background, educational attainment, and municipality of residence.

3.4 Descriptive statistics

Table 1 shows descriptive statistics for three different samples. Column (1) shows descriptive statistics for the full Dutch population for purposes of comparison, while columns (2) and (3) show descriptive statistics for the Nivel sample contacting their GP and the subsample receiving care from an NPMH. The Nivel sample comprises 1,194,388 individuals who first contacted their GP for a mental health–related reason between 2011 and 2021. Of these individuals, 190,105 received treatment by an NPMH at the GP practice within the first 6 months after contact.

The descriptive statistics show several interesting differences and similarities across the three samples. First, individuals seeking care for mental health–related reasons tend to be younger on average, in line with the fact that the onset of many mental health problems occurs at a relatively young age. Additionally, women are significantly likelier to contact their GP and receive treatment from an NPMH. For educational attainment, a slightly different pattern emerges: Individuals with higher education are less likely to contact their GP but likelier to receive treatment from an NPMH. In contrast to what is often found in other countries, migration background does not seem to correlate strongly with care-seeking behavior.

Table 1: Descriptive statistics for general Dutch population, sample contacting their GP for the first time, and sample receiving treatment from an NPMH

	General population[a]	Sample first contacting GP[b]	Sample treated by NPMH[c]
Demographics[d]:			
Age	45.4	43.9	40.4
Male	49.9%	44.2%	38.5%
Dutch born	76.4%	77.7%	77.3%
Education unknown[e]	49.3%	34.9%	24.7%
Education[f]:			
Low	29.6%	34.5%	22.2%
Middle	38.0%	37.1%	44.4%
High	32.4%	28.4%	33.5%
Annual health care (expenditures)[g]**:**			
GP cost	€136.25	€175.60	€173.25
Share using basic MH	1.6%	2.4%	1.8%
Basic MH expenditures	€3.93	€14.32	€11.46
Share using specialized MH	3.8%	7.3%	4.0%
Specialized MH expenditures	€173.14	€457.13	€177.81
Pharmaceuticals	€36.80	€55.69	€35.20
Physical health care	€984.52	€1701.53	€1389.98
Number of individuals	20,633,684	1,194,388	190,105
Survey measures on health[h]**:**			
Self-reported health[i]	2.2	2.3	2.2
High risk of psychological disorder	4.7%	7.7%	9.1%
High risk for anxiety and depression	38.8%	49.0%	53.8%
High risk for loneliness	7.7%	10.6%	11.9%
Number of individuals	365,281	20,957	3,006

(a) Full Dutch population in January 2012. (b) Sample first contacting their Nivel GP for a mental health–related reason between January 2011 and December 2021. (c) Sample referred to an NPMH within first year of contacting their GP. (d) Demographics on January 2012 for the general population or in the year of first contact with GP for remaining two samples. (e) The percentage with an unknown education level in the general population is large because the Dutch education registry started in the 1980s. The education level is unknown for most individuals in earlier cohorts. The difference in the shares with an unknown education level between the general population and the other two samples is driven by age differences. (f) Education level if known. (g) Yearly health care expenditures in 2012 for the general population or in the year before first contact with their GP for the remaining two samples. (h) Survey measures in the 2012 survey wave of the health monitor are available for approximately 2% of the sample. The survey includes only individuals aged 18 or above. (i) Self-reported health is measured on a 5-point Likert scale, with 1 being excellent health.

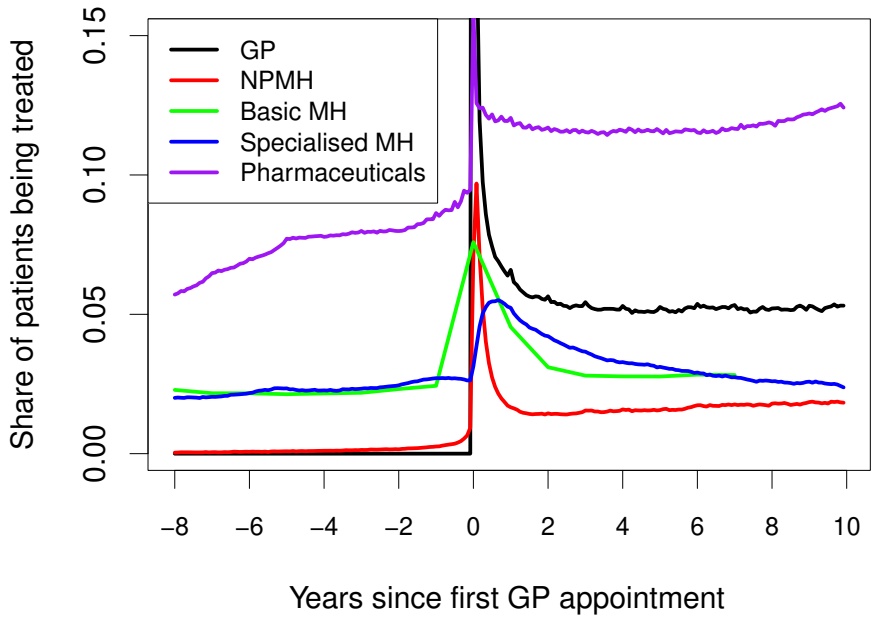
In terms of health care utilization in the year before contacting a GP, the sample first contacting their GP has higher utilization rates across all measures than the general population (whose health care utilization refers to that in 2012). The subsample treated by an NPMH tends to have slightly lower utilization than the sample contacting their GP, which could indicate that treatment by NPMHs is more common among individuals with less severe health issues.

The survey measures of self-reported health largely align with the health care utilization rates: The sample contacting their GP is at higher risk of psychological disorders such as anxiety, depression, and loneliness. Note that the survey outcomes were all measured in 2012. This implies that for the majority of the sample contacting their GP, the survey outcomes were measured years before the individuals sought care.

3.5 Treatment uptake and employment trajectories

To illustrate what happens to patients contacting their GP for a mental health–related reason, Figure 3 shows rates of utilization of the various treatment options relative to when individuals first contacted their GP. By construction, this sample includes only individuals who had not seen their GP for a mental health–related reason in the years

Figure 3: Utilization of treatment types relative to the initial GP appointment



leading up to the first GP appointment, and thus the share seeing their GP jumps to 1 at time zero. The share visiting their GP for mental health–related reasons gradually decreases during the first year, stabilizing at approximately 5% per month.

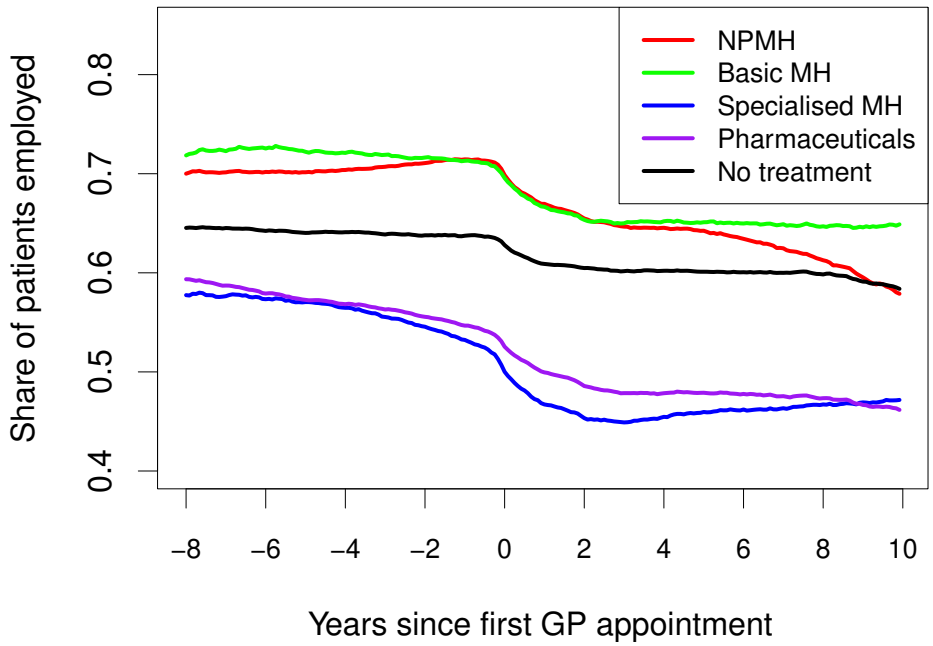
NPMH visits follow a similar but less pronounced pattern, with a peak utilization rate of approximately 10% in the first month after the GP visit. Most NPMH visits take place within the first year after the GP appointment, after which the share receiving this treatment stabilizes at approximately 2%. Note that Figure 3 includes individuals contacting their GP both before and after their GP practice hired an NPMH. In GP practices where talk therapy by NPMHs is available, utilization is thus higher.

Basic and specialized mental health care utilization is lower than NPMH care utilization in the first months after the GP appointment but stabilizes at a higher rate in subsequent years. The slower rise in utilization in the first months after the GP appointment likely reflects the longer waiting times for basic and specialized mental health care than for treatment by an NPMH. The higher rates in later years reflect that the treatment provided in both basic and specialized mental health care facilities tends to be of longer duration than treatment by an NPMH.

Last, a significant share of the sample was already using pharmaceuticals for mental health–related reasons in the years leading up to the first GP appointment. There is a clear spike in the exact month of the GP appointment. Given the long waiting times for basic and specialized mental health care, it is most likely that prescriptions being provided by the GP drive this spike in pharmaceutical utilization. Pharmaceutical utilization stabilizes in subsequent years at a significantly higher level than before the GP appointment, at approximately 12%.

As already briefly mentioned, the type of treatment individuals receive likely depends on a wide range of patients' observable and unobservable characteristics, such as, for example, the severity of their mental health problems. To illustrate these differences, Figure 4 shows the fraction of individuals employed over time across different groups by type of treatment received in the first year after their initial GP appointment. Note that individuals can receive multiple types of treatment and can hence also be in more than

Figure 4: Employment trajectories by type of mental health treatment received in the first year after the initial GP appointment



one of the groups in Figure 4.

We observe substantial divergences in employment rates across the groups receiving the different types of treatment even before the first GP appointment. The individuals who eventually received either NPMH treatment or basic mental health care are substantially likelier to have been employed ex ante than the individuals eventually receiving either no treatment or specialized mental health care and/or pharmaceuticals. Additionally, the drop in the employment rate around the first GP appointment is smaller among patients ultimately treated by an NPMH or in basic mental health care. While this drop in employment around the first appointment should not be interpreted as a causal impact of mental health problems, given the potential for reverse causality and (un)observed confounders, it can be seen as a proxy of the severity of the disorder. While the employment drop among the NPMH group is smaller than that among individuals who ultimately received specialized mental health care, it is still economically meaningful. Approximately 5% of the former individuals lose their jobs around the first GP appointment. This implies that there is indeed scope for mental health treatment to improve the employment outcomes of individuals receiving treatment from an NPMH.

4 Causal impact of talk therapy

As illustrated above, the type of treatment individuals receive correlates with their employment status in the years before they contact their GP. Because this correlation could be due to differences in the severity of the mental health problems they face, a direct comparison of individuals who receive and do not receive treatment from an NPMH is likely to produce biased estimates of the impact of treatment. To resolve this endogeneity issue, I use two different estimation strategies that both exploit the introduction and gradual expansion of the funding available for this type of treatment.

My primary empirical strategy is an IV-style approach taking the GP practice's propensity to refer patients to an NPMH as the instrument. Such differences in referral propensities can be driven by whether an NPMH has already been hired but also by expansions in the number of working hours of the NPMH. This method thus uses all variation in NPMH utilization at the practice level but is predicated on the assumption that a GP practice's propensity to refer patients to an NPMH impacts individuals seeking mental health care solely by increasing their probability of receiving talk therapy. While various checks presented below indicate that this assumption likely holds, unobservable factors could still bias the results.

As a robustness check, I therefore use a regression-discontinuity-in-time (RDiT) strategy that relies solely on the initial hiring of an NPMH at a GP practice. This design compares individuals seeking treatment at a GP practice right before and right after the first NPMH was hired. While this strategy relies on relatively weak assumptions, I can implement it only in the sample of GP practices for which I observe the initial hiring of an NPMH; this restriction shrinks the sample and, correspondingly, the statistical power in estimation.[8] The results from this additional RD estimation strategy, presented in Appendix Section A.3, confirm the main IV results.

[8]As an initial estimation approach, a direct comparison could be implemented between the cohort contacting their GP in the year before the hire and the cohort contacting their GP in the year after the hire. Such a difference-in-difference analysis, however, would not take into account the calendar time difference between these two cohorts. This would result in significant time-varying differences between the cohorts before the initial GP appointment, violating the parallel trends assumption. These results are available upon request.

4.1 Methodology: Differences in the propensity to refer to an NPMH

As the funding available for talk therapy provided by NPMHs increased, GP practices hired NPMHs and expanded their working hours. Variation in the propensity to refer to NPMHs can be driven both by whether the GP practice has hired an NPMH (as the propensity is zero if there is no NPMH) and by increases in the working hours of the employed NPMHs. To exploit this variation, I compute for each GP practice the average monthly probability that a patient seeking care for mental health–related reasons is referred to an NPMH within 12 months of her initial GP appointment.

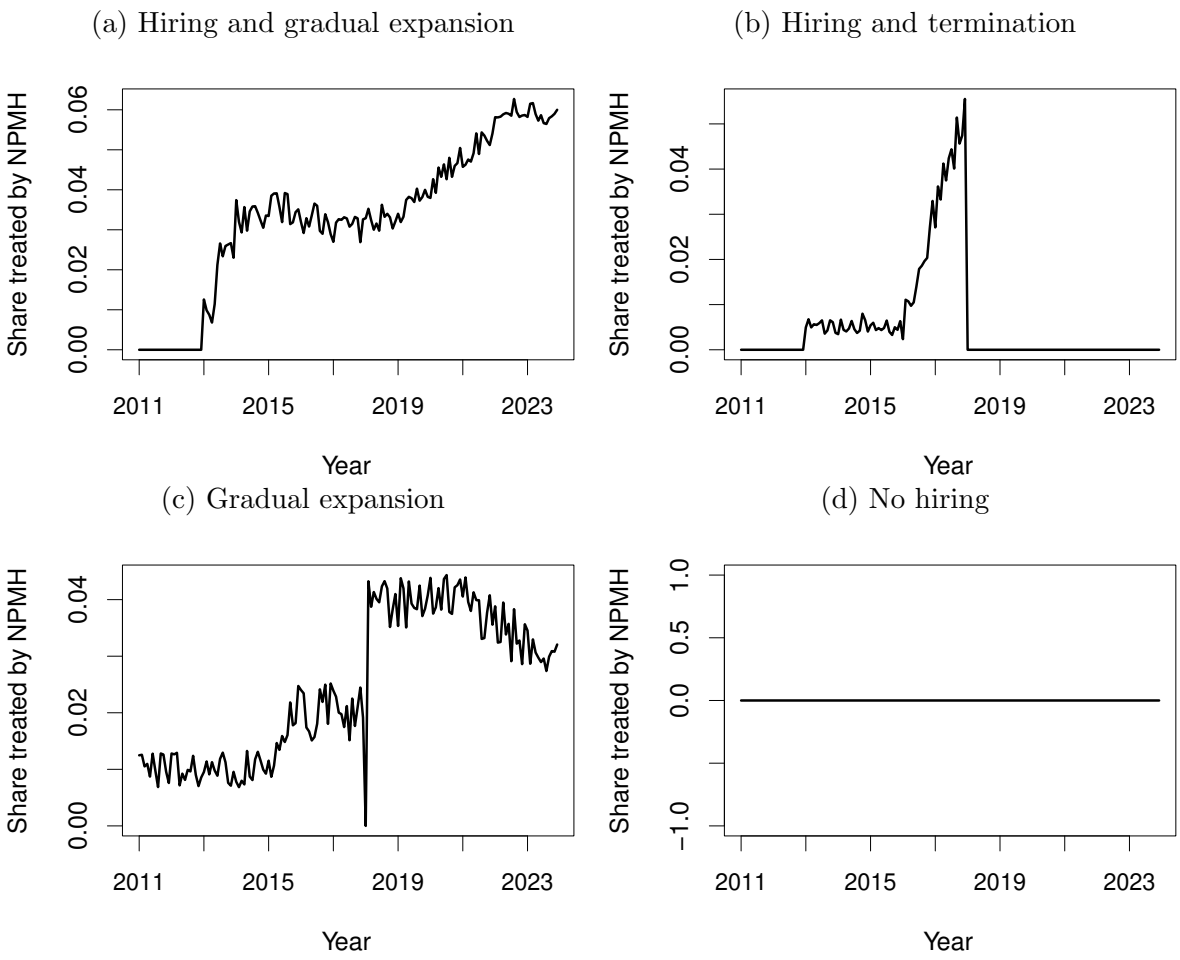
To illustrate, Figure 5 shows these computed probabilities for four different hypothetical GP practices.[9] The practice in panel (a) started referring its patients to an NPMH only in 2013, and its referral rate increased between 2019 and 2020. The practice in panel (b) started referring its patients in 2013 as well, and increased its referrals at a faster rate between 2016 and 2018 as funding increased. However, from 2018 onward, the practice stopped referring patients. For the practice in panel (c), the moment the NPMH was first hired is not observed, as this practice was already referring patients in January 2011. This practice referred patients to the NPMH throughout the study period at a gradually increasing rate until 2020, with a jump in utilization in 2018. From 2020 onward, its referral rate slowly decreased. In contrast to the practices in panels (a), (b) and (c), the practice in panel (d) never hired an NPMH and therefore has a referral rate of zero throughout the observation window.

Given these varying NPMH referral propensities across GP practices, I calculate the patient-level referral propensity as the share of patients receiving treatment by an NPMH within 12 months of their initial GP appointment, within the pool of patients who contacted their GP in the 3 months before their initial GP appointment.[10] These individual propensities are calculated on a leave-one-out basis to ensure that the GP propensity is

[9] Note, due to the privacy regulations of Statistics Netherlands, it is not possible to report results at the individual GP practice level. Hence the plots in Figure 5 show simulated NPMH utilization rates.

[10] As a robustness check, the referral propensity is also calculated as the average probability for patients seeking care in the last month. This yields similar results.

Figure 5: Example of NPMH utilization between 2011 and 2023 at four different hypothetical GP practices



not affected by for example the severity of the mental health problems of the individual for whom the propensity is calculated.

Figure 6 panel (a) shows the average individual propensity to be referred to an NPMH over time, while panel (b) shows the distribution of individual propensities. The expansion of NPMH services occurred primarily between 2011 and 2017, after which the average propensity stabilized at 20%. Panel (b) shows that a large share of individuals seek treatment when the average propensity to refer to an NPMH at their GP practices is zero, as these practices have not hired an NPMH (yet). For practices with an NPMH, the average propensity to refer is between 0 and 40%, indicating that a large share of patients are indeed referred if the option to refer is available.

To estimate the impact of receipt of talk therapy from an NPMH, I estimate the

following two-stage regression:

$$H_{itk} = \alpha_1 + \alpha_2 P_i + \alpha_3 X_{it} + \tau_t + \omega_p + \varepsilon_{it} \quad for \quad -84 \leq k \leq 120$$

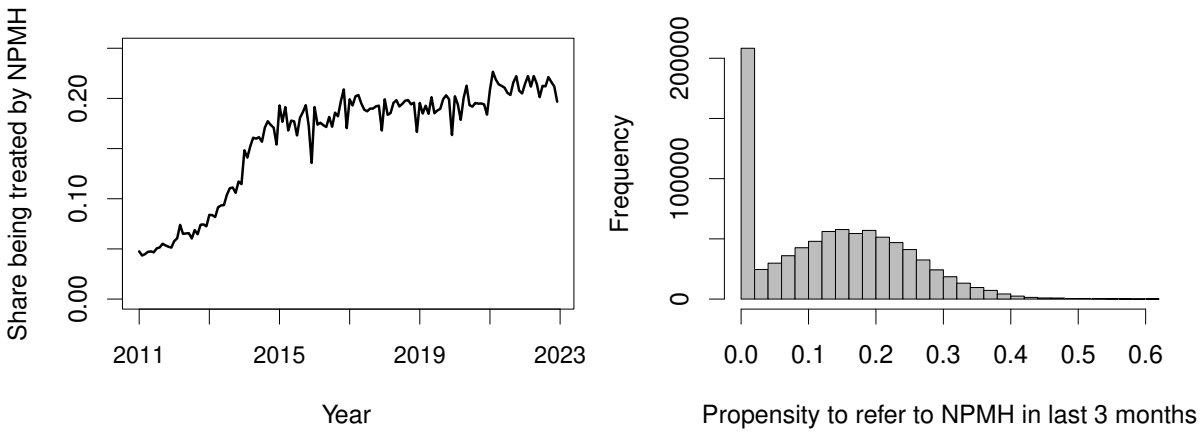
$$E_{itk} = \beta_1 + \beta_2 \hat{H}_{it} + \beta_3 X_{it} + \tau_t + \omega_p + \varepsilon_{it} \quad for \quad -84 \leq k \leq 120$$

where H_{itk} measures health care utilization for individual i seeking treatment at time t and P_{it} is the corresponding propensity to refer to an NPMH at the GP practice of individual i in the 3 months before time t. X_{it} includes various control variables such as the individual's age and gender and the individual's pre–GP appointment health care utilization and labor market status. Finally, τ_t and ω_p are calendar-time and GP practice fixed effects. In the second stage, E_{itk} measures the health and employment outcomes of individual i seeking treatment at time t.[11]

To estimate the dynamic impact of talk therapy provided by NPMHs, I measure the outcomes at month k relative to the moment the individual seeks care, with k corresponding to the period from 84 months before the initial GP appointment to 120 months after it. The outcomes measured in the 84 months before the initial GP appointment are used as placebos. Since the individual had not yet seen her GP at this point, there should be no impact of her GP practice's propensity to refer patients to the NPMH. The estimates for months 0 to 120 show the dynamic impact after the initial GP appointment.

Figure 6: Average GP practice propensity to refer patients to an NPMH between 2011 and 2023 (left) and distribution of GP practice referral propensity

(a) Average referral propensity of GPs (b) Distribution of propensity across patients



[11]See Appendix Table A.4 for a list of all control variables.

IV assumptions

For my estimates of the impact of talk therapy to have a causal interpretation, several crucial assumptions have to be met. First, the instrument (referral propensity) should be relevant, which implies that it should impact a patient's probability of receiving talk therapy. I test this by examining the estimate of α_2 and the F-statistic of the first stage.

Second, the monotonicity assumption should be satisfied, which implies that, for *all* patients, their GP practice's referral propensity (weakly) impacts their probability of receiving care. To test this, I perform heterogeneity analysis on the first-stage results and show that the instrument positively impacts the likelihood of receipt of NPMH care for all subgroups.

Last, the exclusion restriction should hold, such that GP practices' referral propensity impacts only patients through the propensity to receive care. As GP practices with high referral propensities might be of better (or worse) quality than GP practices with low propensities, this assumption could be violated. Similarly, the exclusion restriction could be violated if patients treated by GP practices with high propensities differ from patients at practices with low propensities. To test whether patient or practice differences bias the estimates, I use specifications with and without GP practice fixed effects. These fixed effects should absorb any differences between practices or in patient characteristics.

Even when I rely solely on within-practice variation, the exclusion restriction could still be violated if expansions in NPMH services affect the composition of patients seeking care at the GP practice. GPs could, for example, proactively approach certain patients and offer NPMH services, or patients could seek additional treatment once they learn about the availability of NPMH services. Alternatively, GP practices could expand their NPMH services in response to certain local shocks. To test whether the exclusion restriction holds, I consider placebo outcomes measured before the initial GP appointment, as explained above. Differences in patient composition or NPMH expansions in response to local shocks should show up through significant pretreatment estimates. Additionally, Appendix Table A.5 shows that patient characteristics are not significantly correlated with the instrument, providing further evidence in support of the exclusion restriction.

4.2 First stage: Impact on mental health care treatment uptake

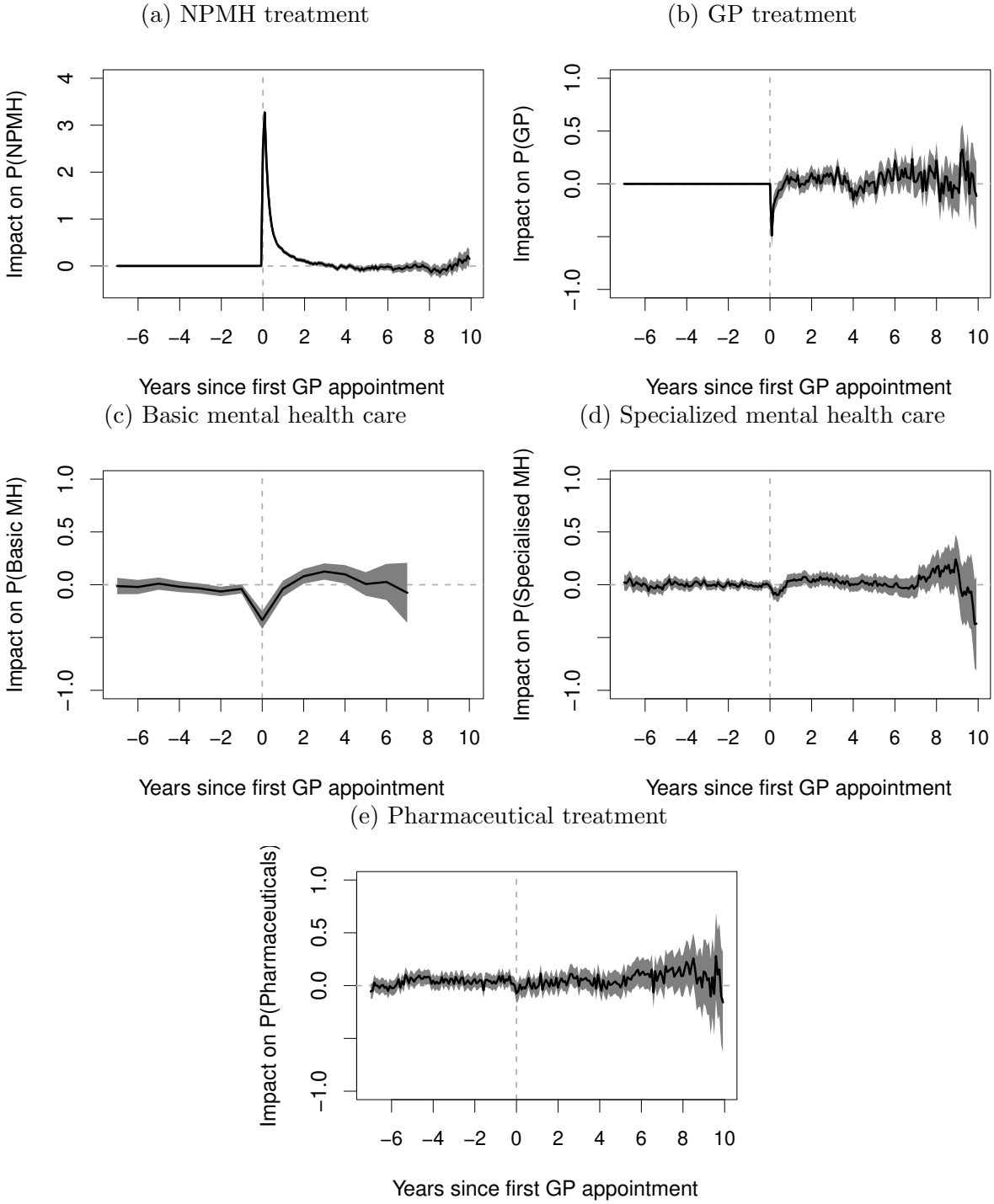
Figure 7 shows the first-stage impact of a 10 pp increase in GP NPMH referral propensity on the probability of receipt of (a) treatment from the NPMH, (b) treatment from the GP, (c) treatment through basic mental health care, (d) treatment through specialized mental health care, and (e) pharmaceutical treatment. The estimates corresponding to months -84 to zero are placebo estimates, as the patients had not yet seen their GP and should thus not be impacted by the GP's propensity to refer to an NPMH. Reassuringly, all the placebo estimates are not significantly different from zero.

The estimates corresponding to months 0 to 120 show the impact of the GP referral propensity on health care utilization in the 10 years after the initial GP appointment. Panel (a) shows that an increase in NPMH referral propensity significantly increases the probability that a patient indeed receives treatment from the NPMH. Uptake of talk therapy spikes in the month of the initial GP appointment but quickly attenuates in the months that follow. This is in line with the relatively low intensity and duration of the treatment provided by the NPMH. Panels (b) and (c) show that the increase in NPMH treatment is mirrored by reductions in the probability of patients visiting the GP or receiving basic mental health care. These reductions are, however, of much smaller magnitude than the increase in NPMH utilization. There are no impacts on specialized mental health care (panel (d)) or the use of pharmaceuticals (panel (e)).

To further assess the magnitude and dynamic impact of NPMH talk therapy, I tabulate in Table 2 the cumulative effects on the same health care outcomes (rows) across varying time windows (columns). The results for quantities of care instead of the use of specific types of care are reported in Appendix Table A.8. All estimates correspond to the impact of a 10 pp increase in GP NPMH referral propensity. Given the large number of outcomes considered, I correct all p-values for multiple hypothesis testing using the Benjamini & Hochberg (1995) procedure.

Similarly to Figure 7 above, the first column shows placebo estimates of the effect on health care utilization in the 12 months before the initial GP appointment. As the patients had not yet seen their GP, there should be no impact of GPs' referral propensity.

Figure 7: Impact of a 10 pp increase in GP NPMH referral propensity on health care utilization in the 7 years prior to 10 years after the initial GP appointment



Reassuringly, the estimates for basic or specialized mental health care utilization and the use of pharmaceuticals are again not significantly different from zero and small in magnitude. This implies that patients at GP practices with high propensities are similar to those at GP practices with low propensities, providing further evidence in support of

Table 2: Differences in treatment uptake rates in the 12 months prior to the GP appointment and 12, 24, 48 and 96 months after the GP appointment

	12 pre-months	12 months	24 months	48 months	96 months
P(NPMH visit)		5.01%**	4.95%**	4.35%**	3.46%**
		(0.05)	(0.06)	(0.06)	(0.07)
P(GP visit>1)		0.18%*	0.29%**	0.28%**	0.25%**
		(0.08)	(0.08)	(0.07)	(0.07)
P(Basic MH)	-0.04%	-0.33%**	-0.35%**	-0.28%**	-0.27%**
	(0.02)	(0.04)	(0.04)	(0.05)	(0.05)
P(Specialized MH)	0.02%	-0.10%	-0.09%	-0.05%	-0.11%
	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)
P(Pharmaceuticals)	0.06%	-0.03%	-0.03%	0.01%	-0.06%
	(0.05)	(0.06)	(0.06)	(0.07)	(0.07)

Standard errors shown in parentheses; *significant at 5%; **significant at 1%; significance levels reported after correction for multiple hypothesis testing with the Benjamini & Hochberg (1995) procedure.

the exclusion restriction.

The estimated effects on health care utilization in the first 12, 24, 48, and 96 months after the initial GP are in line with the dynamic estimates in Figure 7. Patients seeking care at GP practices with high practice propensities are likelier to have received talk therapy from an NPMH (row 1). A 10 pp increase in the practice's propensity to refer increases an individual's probability of receiving treatment by 5 pp. The cumulative impacts exhibit a decreasing pattern as the time window expands, indicating that patients facing lower GP propensities partly catch up with those facing higher GP propensities. However, even when I consider a time window of 8 years after the initial GP appointment, patients facing higher GP propensities are still significantly likelier to have received talk therapy from the NPMH. To facilitate interpretation of the magnitude of the increased take-up, the RDiT analysis in Appendix Section A.3 shows that the probability of receipt of talk therapy from an NPMH increases by approximately 6 pp once an NPMH is first hired at a GP practice.

Row 2 shows that the decrease in GP visits in the first months after the initial appointment, as seen in Figure 7 panel (b), is offset by an increase in GP visits in later months. All the GP visit coefficients are positive but smaller in magnitude than the NPMH coefficients. There is a significant reduction in utilization of basic mental health

care across all the time windows considered. The magnitude of this reduction is, however, smaller by a factor of 10 than the increase in NPMH utilization. When I consider annual expenditures for basic mental health care, this lower utilization translates to a reduction of approximately €2, indicating that the patients who no longer received basic mental health care would have received only a very limited amount of basic mental health care (see Appendix Table A.8). The estimates for specialized mental health care and the use of mental health–related pharmaceuticals are small and not significant.

The IV estimates prove robust across a range of alternative specifications, as shown in Appendix Table A.7. The specification using only the 2011–2017 time window, when NPMH funding experienced the largest expansions, yields similar estimates, as does the one excluding observations with zero propensity to refer to the NPMH. Exclusion of the GP fixed effects does result in larger estimates for basic and specialized mental health care. This implies that there are differences between GP practices with high and low NPMH referral propensities. Estimates without GP fixed effects, such as those presented in Dros et al. (2024), should thus not be interpreted as causal. Last, the estimates exploiting the first hiring of NPMHs at GP practices in the RDiT framework point to similar impacts, as shown in Appendix Section A.3.

To summarize, the first-stage estimates show that the introduction of talk therapy at GP practices significantly increased uptake of the new treatment, with limited substitution away from the other treatment options. Further analyses looking at various combinations of mental health care as outcomes show that the increase in talk therapy uptake is driven largely by individuals who would not have received any care in the absence of the NPMH (80% of all patients). The remaining 20% of the patients being treated by the NPMH receive talk therapy on top of other forms of mental health care (see Appendix Table A.6). In short, for most patients, talk therapy at GP practices complements existing treatment options. The following subsection analyzes whether this increased treatment uptake reduced or widened existing disparities in mental health care utilization.

Heterogeneity in the impact of talk therapy

Disparities in mental health care utilization have been reported in various countries (Sentell et al., 2007; Cook et al., 2017). Whereas the literature has focused solely on differences in mental health care uptake, in Appendix Section B, I estimate these disparities in the Netherlands combining both mental health care utilization and mental health care *need*. I show that incorporating the latter variable, measured by self-reported mental health, is important as it affects both the sign and magnitude of the estimated disparities. Using the information on mental health care utilization and need, I find that men, older individuals and people with either a migration background or lower educational attainment use disproportionately little mental health care compared to their counterparts.

To determine whether the introduction of talk therapy at GP practices reduced some of these disparities, consider the results in Table 3, which shows the impacts across various subgroups. These estimates are obtained through a split-sample approach. Column (1) shows the impact of a 10 pp increase in the GP NPMH referral propensity, and column (2) shows the impact on receipt of any mental health care treatment in the first 12 months after the initial GP appointment. The estimates indicate that younger individuals, women, those who are Dutch born, and individuals with middle or high educational attainment were likelier to receive talk therapy and, correspondingly, to receive any mental health care. These groups that saw the largest increase in treatment uptake are the ones that already used relatively more mental health care. The introduction of talk therapy thus exacerbated mental health care disparities.

Table 3: Impact of a 10 pp increase in GP NPMH referral propensity on mental health care utilization in the first 12 months after the initial GP appointment for various population subgroups

	P(NPMH)	P(Any MH treatment)	# observations
Full sample	5.01%**	3.69%**	870.550
	(0.05)	(0.07)	
Age $\leq$ 45	5.81%**	4.08%**	441.484
	(0.08)	(0.10)	
45 $\leq$ Age $\leq$ 65	5.40%**	4.22%**	246.202
	(0.11)	(0.14)	
Age >65	2.38%**	1.94%**	252.989
	(0.09)	(0.11)	
Male	4.31%**	3.18%**	385.921
	(0.08)	(0.10)	
Female	5.59%**	4.11%**	484.629
	(0.08)	(0.10)	
Dutch born	5.03%**	3.81%**	678.535
	(0.06)	(0.08)	
Non-Dutch born	4.92%**	3.23%**	192.015
	(0.11)	(0.15)	
Education unknown	3.79%**	3.04%**	314.333
	(0.08)	(0.10)	
Education low	3.74%**	2.79%**	187.176
	(0.10)	(0.14)	
Education middle	6.60%**	4.72%**	211.504
	(0.12)	(0.16)	
Education high	6.43%**	4.35%**	157.537
	(0.15)	(0.19)	

Standard errors shown in parentheses; *significant at 5%; **significant at 1%, significance levels reported after correction for multiple hypothesis testing with the Benjamini & Hochberg (1995) procedure.

4.3 Second stage: Impact of talk therapy on health and labor market outcomes

The previous subsection showed that the introduction of talk therapy at GP practices significantly increased patients' probability of receiving it. Did this increase in treatment uptake also improve the health and labor market outcomes of the individuals receiving more care? To answer this question, I estimate the following slightly adjusted two-stage regression:

$$\begin{aligned} H_{it} &= \alpha_1 + \alpha_2 P_i + \alpha_3 X_{it} + \tau_t + \omega_p + \varepsilon_{it} \\ E_{itk} &= \beta_1 + \beta_2 \hat{H}_{it} + \beta_3 X_{it} + \tau_t + \omega_p + \varepsilon_{it} \quad for \quad -96 \leq k \leq 120 \end{aligned}$$

Since most of the talk therapy takes place within the first 12 months after the initial GP appointment, I use an indicator for receipt of any talk therapy within those 12 months as my outcome in the first stage (H_{it}) and as the main independent variable in the second stage ($\hat{H}_{it}$). As shown in Table 2, a 10 pp increase in GP referral propensity increases the probability of talk therapy uptake in the subsequent 12 months by 5 pp, corresponding to α_2. The F-statistic of this first stage estimate is 8491, implying high relevance of the instrument.

Impact of talk therapy on health outcomes

There are no full-population measures of health available in the Dutch administrative data. To estimate the impact of talk therapy on health, I therefore use self-reported health measures obtained from the national health monitor and physician-assessed symptom severity in specialized mental health care. Both of these measures are only available for a small share of the population and are measured infrequently. Hence, I look at health measures averaged over the 4 years before the initial GP appointment (as a placebo outcome) and the 4 years after the initial GP appointment.[12] [13]

[12] Using shorter (or longer) time windows results in less (more) power but yields similar results.

[13] Given the much smaller sample sizes for health outcomes, results using the alternative RDiT strategy are not reported in Appendix Section A.3. All RDiT estimates are insignificant but have very wide confidence intervals.

While self-assessed health measures are available for a random subsample, physician-assessed health measures are available only for individuals who receive treatment in specialized mental health care. For these individuals, the healthcare provider makes an assessment of the severity of their symptoms at the start of treatment and (infrequently) while treatment is underway. As this measure is only available for individuals who receive treatment in specialized mental health care, it is only available for individuals with more severe mental health problems. Additionally, differences in physician-assessed health may be biased if talk therapy significantly affects the probability of receiving specialized mental health care (and thus the likelihood of observing these health measures). However, as shown in the previous section, there is no significant impact of talk therapy on the probability of receiving specialized mental healthcare, alleviating some of these concerns.

Table 4 shows the impact of talk therapy on self-assessed and physician-assessed (mental) health.[14] For ease of interpretation, all health measures are normalized to have a mean of zero and a standard deviation of one, with positive values corresponding to better (mental) health. For both self-assessed and physician-assessed measures, there is no significant impact of talk therapy on average health in the 4 years before the initial GP appointment, further increasing the credibility of the IV estimation strategy.

The second column of Table 4 shows the impact of receiving talk therapy on health measures in the first 4 years after the initial GP appointment. For the self-assessed health measures reported in the national health monitor, there are no significant improvements in general or mental health. Additionally, measures corresponding to the risk of anxiety or depressive disorders and loneliness are unaffected by talk therapy. The lower half of Table 4 presents the impact of talk therapy on the severity of symptoms at the beginning and end of specialized mental health care, as well as the worst severity assessed during the treatment trajectory. All estimates are small and not significantly different from zero, indicating that talk therapy does not reduce symptom severity.

It is important to re-emphasize that the health measures are only available for a relatively small subsample of the full population. The resulting power of the analysis

[14] Appendix Table A.9 shows the reduced form estimates of a 10 percentage point increase in GP propensity to refer to the NPMH on these health outcomes.

Table 4: Impact of talk therapy on self-assessed and physician-assessed health in the 48 months prior to the GP appointment and 48 months after the initial GP appointment

	48 pre-months	48 post-months
Self-assessed health[a]:		
General health	-0.01	0.10
	(0.19)	(0.20)
Mental health	-0.10	0.05
	(0.12)	(0.15)
Risk of anxiety disorder or depression	-0.24	-0.24
	(0.20)	(0.22)
Loneliness	-0.27	0.09
	(0.21)	(0.22)
Sample size	24,394	19,301
Physician-assessed health[b]:		
At start of treatment	-0.10	-0.12
	(0.11)	(0.10)
At end of treatment	-0.03	-0.04
	(0.11)	(0.10)
Worst assessed	0.09	-0.01
	(0.10)	(0.10
Sample size	95,865	91,045

Standard errors shown in parentheses; *significant at 5%; **significant at 1%; significance levels reported after correction for multiple hypothesis testing with the Benjamini & Hochberg (1995) procedure. [a] Measures of self-assessed health reported in the national health monitor, [b] Physician-assessed health as assessed during treatment in specialized mental health care

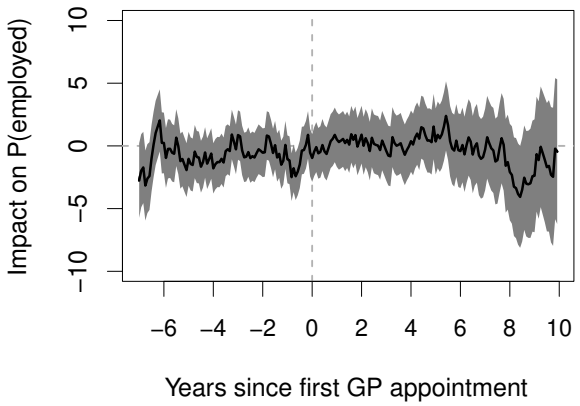
is insufficient to estimate impacts on health in the first year following the initial GP appointment, when most of the talk therapy takes place. The analysis thus shows that there are no long-run improvements in health attributable to talk therapy. There could potentially still be short-run improvements in (mental) health.

Impact of talk therapy on employment outcomes

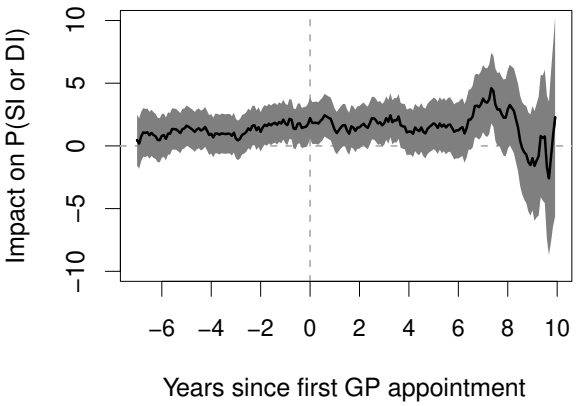
Finally, I turn to the impact of talk therapy on employment outcomes. The results discussed below show the impact of receiving talk therapy within the first 12 months after the initial GP appointment on a range of labor market outcomes. For completeness, Appendix Figure A.1 also shows the reduced form labor market impacts, i.e. the impact of a 10 pp increase in GP propensity to refer to the NPMH on labor market outcomes.

Figure 8: IV: Impact of receiving talk therapy in the first 12 months after the initial GP appointment on employment, DI/SI benefit receipt and earnings measured starting 7 years before the initial GP appointment until 10 years after the initial GP appointment.

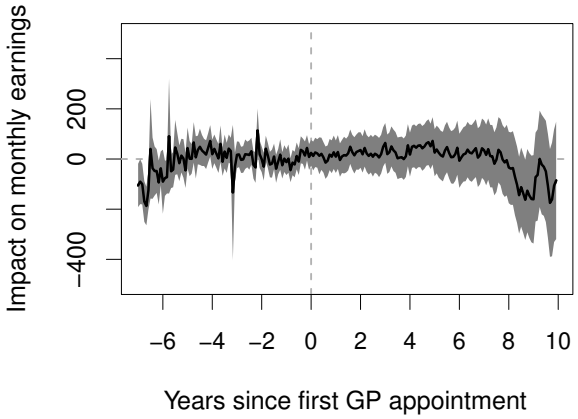
(a) Employment



(b) DI/SI benefit receipt



(c) Earnings



Reassuringly, the placebo estimates examining labor market outcomes before the initial GP appointment are small and insignificantly different from zero. This implies that the pre-treatment labor market outcomes of patients facing a high GP propensity to refer are similar to those of patients facing a low GP propensity to refer.

However, for all three labor market outcomes, there are also no significant improvements as a result of receiving talk therapy in the 10 years after the initial GP appointment. Given the sample size, these IV estimates make it possible to reject economically meaningful improvements in labor market outcomes as a result of the introduction of talk therapy at GP practices. To summarize the impact of talk therapy on the intensive and extensive margin of employment over an extended time window, I estimate the impact on average monthly labor earnings in the first 10 years after the initial GP appointment. As

expected, the estimate is small (€9 on a mean of €1,600) and insignificant (€31 standard error).

The absence of labor market improvements contradicts recent research on the NHS Talking Therapies program, finding significant employment gains from talk therapy Rzepnicka et al. (2025). To determine whether this contradiction is caused by differences in the setup of the Dutch and UK programs, or by methodological biases, I apply the methodology of Rzepnicka et al. (2025) to the Dutch context by comparing patients who only receive 1 session of talk therapy from the NPMH to those who receive more than 1 session of talk therapy. Using this methodology, I do find significant (but negative) impacts of talk therapy on labor market outcomes (see Appendix Figure A.5). Given that this analysis is performed on the same sample as the main IV analysis in this paper, these significant estimates cannot reflect a causal impact but rather indicate a methodological bias. The pre-trend estimates (especially for earnings) also indicate that individuals who choose to drop out of treatment after 1 session are on a different trajectory than those who decide to continue treatment.

5 Conclusion

Mental health disorders impose a substantial (economic) burden on society, and many people with these disorders do not receive any treatment (OECD, 2014; NIMH, 2024). In an attempt to increase treatment uptake and reduce the negative burden of mental health, various countries and US states introduced easily accessible forms of mental health care in the 2000s and 2010s. Examples of these programs are the Improving Access to Psychological Therapies (United Kingdom, now called NHS Talking Therapies), Better Access Initiative (Australia), Ontario Structured Psychotherapy Program (Canada), Prompt Mental Health Care (Norway), and Mental Health in Primary Care (Netherlands), with community mental health centers or organizations such as Sondermind or Open Path Collective providing similar types of care in the US.[15]

These programs often provide relatively low-intensity treatment (talk therapy) at low

[15]See Appendix Table A.1 for a comparison of the main characteristics of the various programs.

or zero cost to patients on a large scale. In the Netherlands, the UK, and Australia, for example, 1 to 3% of the entire population (0.5 to 1.3 million individuals) receives mental health care through these programs each year.

While the medical literature has shown that the type of treatment these programs provide effectively reduces the health burden of mental disorders (Hofmann et al., 2012), causal evidence on the effectiveness of the national roll-out of these programs is limited. In particular, it is not yet known how these programs interact with existing treatment options and whether they can reduce existing disparities in mental health care utilization. Additionally, evidence on whether or not this type of treatment is also effective in reducing the (economic) burden of mental health issues is also lacking.

In this paper, I fill this gap in the literature by examining a reform in the Netherlands that introduced an easily accessible form of talk therapy provided by nurse practitioners at General Practitioner (GP) practices. I first determine whether this introduction has increased the uptake of mental health treatment for the full population and, if so, whether this expansion of treatment options has reduced or increased existing disparities in healthcare utilization. Next, I examine whether patient health and labor market outcomes have also improved as a result of the increased treatment uptake.

By exploiting the introduction of funding availability through both a Regression-Discontinuity-in-Time and Instrumental Variable framework, I find a significant increase in the uptake of talk therapy among patients seeking care. The uptake of already existing forms of mental health care is affected to a much smaller extent, implying that for most patients, talk therapy provided at GP practices acts as a complement and not a substitute for these other forms of treatment. However, there are also no long-term reductions in mental health care utilization, pointing to limited preventive healthcare gains of talk therapy.

Next, I examine whether the increased treatment uptake reduced or widened the existing disparities in healthcare utilization. Whereas previous literature has estimated these disparities by examining only utilization rates, I combine measures of utilization and need for care. Controlling for the need for care substantially impacts both the

magnitude and the sign of the estimated disparities. Combining measures of healthcare utilization and need shows that older individuals, males, and those with either a migration background or lower educational attainment use disproportionately little mental health care.

With these existing disparities in mind, I estimate heterogeneous impacts of the introduction of talk therapy at GP practices and find that younger individuals, females, and those without a migration background or higher educational attainment experience the most significant increases in treatment uptake. These groups were already relatively likelier to use existing forms of mental health care. The introduction of talk therapy at GP practices has thus increased existing disparities in mental health care utilization.

Finally, I assess the impact of talk therapy on the health and labor market outcomes of patients. Despite the significant increase in uptake of talk therapy, there are no improvements in a wide range of patient outcomes. Firstly, various measures of self-assessed and physician-assessed (mental) health do not significantly improve in the long-run.[16] Additionally, employment, labor earnings, and the receipt of various social benefits are unaffected in the short and long run (up to 10 years after treatment).

The absence of impacts of the program on already-existing types of mental health care and on health and labor market outcomes contrasts with prior literature examining the impact of similar programs in the Netherlands and the UK. Specifically, Dros et al. (2024) find that talk therapy reduces referrals to basic/specialized mental health care in the Netherlands. I show that these associations are likely driven by differences between GP practices and do not reflect the causal impact of talk therapy. Similarly, Oparina et al. (2024) and Rzepnicka et al. (2025) use waitlisted patients and individuals who drop out of treatment as control groups and find that patients of the equivalent UK program experience higher recovery rates and improved labor market outcomes. While the UK program could potentially be more effective, these estimates could also reflect differences between individuals who do and do not drop out of treatment thereby leading to biased estimates. To test this, I apply the methodology of Rzepnicka et al. (2025)

[16] Due to limited data availability, impacts on health in the short-run cannot be estimated with enough power.

to the Dutch setting, comparing patients who drop out after one session to those who continue treatment and show that this indeed yields biased estimates (Appendix Figure A.5).

Lastly, the null results on labor market outcomes are also in contrast with prior research on the causal impact of mental health treatment. Biasi et al. (2021); Shapiro (2022) and Prudon (2025) all find significant impacts on labor market outcomes due to treatment against bipolar disorder, the use of antidepressants, and waiting times for mental health treatment. One potential explanation for this discrepancy is that all three papers examine treatment that incorporates (or is strictly limited to) the use of pharmaceuticals. In contrast, treatment in this paper (talk therapy) largely excludes pharmaceuticals. Additionally, Biasi et al. (2021) and Prudon (2025) focus on relatively severe mental health problems. Patients receiving talk therapy at GP practices arguably suffer from less severe mental health problems. However, it is essential to note that these patients still experience a persistent five percentage point drop in their probability of employment around the moment they seek treatment. There is thus scope for treatment to have a positive labor market impact.

A second potential reason for the lack of health and labor market impacts could be the intensity of treatment. While the introduction of the talk therapy option does increase treatment take-up by approximately 6 percentage points, individuals, on average, only receive about five 30-minute sessions in the first year of treatment. While this is similar to the average treatment in for example the NHS Talking Therapies program in the UK, the treatment is arguably of a very low intensity (see Appendix Table A.1 for a comparison of the treatment intensity in various programs). As a point of comparison, individuals being treated in specialized mental health care treatment receive approximately eight times as many sessions. Further research is required to determine whether talk therapy at a higher intensity could improve health and labor market outcomes.

To summarize, the aim of introducing talk therapy provided by NPMHs at GP practices was to increase mental health treatment uptake while reducing existing health care disparities and alleviating pressure on already-existing treatment options. The increased

uptake of treatment should improve the long-run (socioeconomic) outcomes of individuals struggling with mental health problems. These aims should be accomplished at relatively low monetary cost. The results in this paper highlight that although treatment uptake was significantly increased, the other policy goals were not achieved. Existing health care disparities are exacerbated as subgroups of the population that already used relatively much care before the introduction are most likely to adopt the new form of treatment. Additionally, the utilization of already-existing treatment options is largely unaffected and the health and socio-economic outcomes of patients has not improved.

In terms of cost, the increased treatment uptake was achieved at a relatively low per-patient cost of €500, compared to an average per-patient cost of €5000 in specialized health care. However, given the large number of individuals receiving treatment, the total cost of the program sums to approximately €250 million per year, on a mental health care budget of €5 billion. Additionally, the program employs 3,000 full-time-equivalent nurse practitioners, imposing significant human-capital costs on a sector facing labor shortages.

While the increased treatment uptake may have improved short-run patient wellbeing, the absence of an impact on future health care utilization and lack of health and labor market improvement indicates that more or different interventions are needed to reduce health care disparities and improve the long-term outcomes of individuals suffering from mental health problems.

References

Alonso, J., Liu, Z., Evans-Lacko, S., Sadikova, E., Sampson, N., Chatterji, S., ... others (2018). Treatment gap for anxiety disorders is global: Results of the world mental health surveys in 21 countries. *Depression and anxiety*, *35*(3), 195–208.

Bastiampillai, T. J., Allison, S., Harford, P., Perry, S. W., & Wong, M.-L. (2019). Has the uk improving access to psychological therapies programme and rising antidepressant use had a public health impact? *The Lancet Psychiatry*, *6*(3), e8–e9.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289–300.

Biasi, B., Dahl, M. S., & Moser, P. (2021). Career effects of mental health. *NBER working paper, 29031*.

Bruun, L. S., Hróbjartsson, A., Gluud, C., Kongerslev, M. T., Simonsen, E., Storebø, O. J., et al. (2022). Control interventions in randomised trials among people with mental health disorders. *Cochrane Database of Systematic Reviews*(4).

Clark, D. M., Canvin, L., Green, J., Layard, R., Pilling, S., & Janecka, M. (2018). Transparency about the outcomes of mental health services (iapt approach): an analysis of public data. *The Lancet*, *391*(10121), 679–686.

Cook, B. L., Trinh, N.-H., Li, Z., Hou, S. S.-Y., & Progovac, A. M. (2017). Trends in racial-ethnic disparities in access to mental health care, 2004–2012. *Psychiatric Services*, *68*(1), 9–16.

Dros, J. T., van Dijk, C. E., Böcker, K. B., Bruins Slot, L. C., Verheij, R. A., Meijboom, B. R., ... Bos, I. (2024). Healthcare utilization patterns of individuals with depression after national policy to increase the mental health workforce in primary care: a data linkage study. *BMC primary care*, *25*(1), 1–10.

government, A. (2025). *Better access initiative.* https://www.health.gov.au/our-work/better-access-initiative. ([Accessed 24-09-2025])

Hauenstein, E. J., Petterson, S., Merwin, E., Rovnyak, V., Heise, B., & Wagner, D. (2006). Rurality, gender, and mental health treatment. *Family & Community Health*, *29*(3), 169–185.

Health, O. (2025). *Ontario structured psychotherapy program.* https://www.ontariohealth.ca/clinical/mental-health-addictions/structured-psychotherapy. ([Accessed 24-09-2025])

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive therapy and research*, *36*, 427–440.

Jones, P. B. (2013). Adult mental health disorders and their age at onset. *The British Journal of Psychiatry*, *202*(s54), s5–s10.

Jorm, A. F. (2018). Australia's 'better access' scheme: Has it had an impact on population mental health? *Australian & New Zealand Journal of Psychiatry*, *52*(11), 1057–1062.

Kohn, R., Saxena, S., Levav, I., & Saraceno, B. (2004). The treatment gap in mental health care. *Bulletin of the World health Organization*, *82*(11), 858–866.

Lancet. (2020). Mental health matters. *The Lancet. Global Health*, *8*(11), e1352.

NHS. (2024). *Nhs talking therapies.* https://www.england.nhs.uk/mental-health/adults/nhs-talking-therapies/. ([Accessed 24-09-2025])

NIMH. (2024, September). *Mental illness.* Retrieved from https://www.nimh.nih.gov/health/statistics/mental-illness

OECD. (2014). Making mental health count. *OECD Health Policy Studies*, 244.

Oparina, E., Krekel, C., & Srisuma, S. (2024). *Talking therapy: Impacts of a nationwide mental health service in england* (Tech. Rep.). IZA Discussion Papers.

POH-GGZ, L. (2020). *Functie- en competentieprofiel 'praktijkondersteuner huisarts ggz' 2020.* https://poh-ggz.nl/storage/documents/2042/functie-en-competentieprofiel-poh-ggz.pdf. ([Accessed 24-09-2025])

Prudon, R. (2025). Is delayed mental health treatment detrimental to employment? *Review of Economics and Statistics*, Forthcoming.

Rzepnicka, K., Sharland, E., Rossa, M., Dolby, T., Oparina, E., Saunders, R., ... Nafilyan, V. (2025). The effect of adult psychological therapies on employment and earnings: Evidence from england. *medRxiv*, 2025–01.

Sæther, S. M. M., Knapstad, M., Grey, N., & Smith, O. R. (2019). Twelve months post-treatment results from the norwegian version of improving access to psychological therapies. *Frontiers in Psychology*, *10*, 2303.

Sentell, T., Shumway, M., & Snowden, L. (2007). Access to mental health treatment by english language proficiency and race/ethnicity. *Journal of General Internal Medicine*, *22*(2), 289–293.

Shapiro, B. T. (2022). Promoting wellness or waste? Evidence from antidepressant advertising. *American Economic Journal: Microeconomics*, *14*(2), 439–77.

WHO. (2022, June). *Factsheet mental disorders.* Retrieved from https://www.who.int/news-room/fact-sheets/detail/mental-disorders

A Appendix

A.1 Additional descriptive results

Table A.2: Names of the various Statistics Netherlands datasets used including a brief description

Dataset name	Description
Healthcare:	
Nivel	GP and NPMH treatment information
GGZDBCGeleverdzorgprofiel	Treatment activity dates, treatment minutes and provider type
GGZDBCTrajecten	First moment of contact and diagnoses
ZVWZorgkostentab	Annual health care expenditures
MedicijnTab[a]	Use of pharmaceuticals
Demographic characteristics:	
GBAPersoon	Gender, age and migration status
HoogsteoplTab	Education level
GBAAdresobject and VSLGWB	Municipality of residence
Employment outcomes:	
Secmbus	Indicators for various employment status
Polis	Earnings and working hours

Note: [a] MedicijnTab contains yearly indicators of pharmaceutical utilization by ATC-4 code. For this paper, Statistics Netherlands has provided a custom version of MedicijnTab which contains exact dates at which prescriptions are filled for a subset of ATC-4 codes related to mental health disorders.

Table A.3: Construction of mental health care expenditures and physical health care expenditures based on expenditure categories used by Statistics Netherlands

Mental health care	Non-mental health care	Pharmaceuticals
First-line psychological health care	General practitioner	Pharmacy
Mental health care	Hospital health care	
Basic-mental health care	Paramedical health care	
Specialist mental health care	Nursing without stay	
Geriatric rehabilitation health care		

Note: Several expenditure categories, such as health care abroad and other costs, as used by Statistics Netherlands are excluded from all categories as it is not clear to which category they belong.

Table A.1: Comparison of the characteristics of the various programs offering easily accessible forms of talk therapy

Characteristics	**NPMH**	**NHSTT**	**BAI**	**OSP**	**PMHC**
Country/State	Netherlands	UK	Australia	Ontario, Canada	Norway
Full name	Nurse Practitioners Mental Health at GP practices	NHS Talking Therapies, previously called IAPT	Better Access Initiative	Ontario Structured Psychotherapy Program	Prompt Mental Health Care
Treatment type	Provide diagnostics, mental coaching and CBT-inspired techniques	Face-to-face, remote, group or individual forms of talk therapy	Rebates for existing mental health care through Medicare	Cognitive behavioral therapy and related services	Cognitive behavioral therapy
Treatment provider	Nurse practitioners specialised in mental health care	Psychological wellbeing practitioners and therapist	GPs, psychologists, social workers, occupational therapists and other medical practitioners	Coach or therapist	Multidisciplinary team with a psychologist
Maximum treatment intensity	No maximum	No maximum	6 sessions per treatment course and 10 individual plus 10 group sessions per calender year	No maximum	No maximum
Average treatment intensity	6 sessions of 30 minutes	8 sessions of 30-50 minutes	5 sessions	Not reported	5 sessions
Pharmaceuticals	No, only provided through GP	No, only provided through GP	Yes, but cost covered through separate program (PBS)	No, provided through GP	No, provided through GP
Patient population	Full population	Adults with depression and anxiety	Individuals diagnosed with mental disorder	Adults with depression and anxiety-related concerns	Full population
Cost of care	Covered by mandatory health insurance	Free of charge	Co-payment for some sessions (65%) of on average €50	Free of charge	Free of charge

Data sources: (1) Netherlands: POH-GGZ (2020), (2) UK: NHS (2024), (3) Australia: government (2025), (4) Ontario, Canada: Health (2025)

A.2 Additional results IV analysis

Table A.4: All individual control variables X_i used in the IV specification

Control variable	Values
Demographics:	
Gender	Male or Female
Age	Dummies for 5-year age groups
Migration status	Native, first or second generation
Education level	Low, middle, high or unknown
Calendar-month of first contact	Year-month FE
GP practice FE	
Pre-treatment employment[a]:	
Employment	Indicator for being employed
UI	Indicator for receiving UI
Social assistance	Indicator for receiving social assistance
Sickness/DI	Indicator for receiving sickness/DI benefits

(a) Pre-treatment employment outcomes are measured 12 months prior to initial GP appointment, or 12 months prior to outcome in case of placebo outcomes

Table A.5: Balance of individual characteristics on instrument

Characteristic	Impact of 10 p.p. GP practice increase
Age	-0.89
	(2.64)
Male	-0.07%
	(0.08)
Native	-0.04%
	(0.06)
Education unknown	0.12%
	(0.06)
Education low	-0.06%
	(0.08)
Education middle	0.10%
	(0.09)
Education high	-0.04%
	(0.08)

Standard errors shown in parentheses; *significant at a 5% significance level; **significant at a 1% significance level, significance level reported after correcting for multiple hypothesis testing using the Benjamini & Hochberg (1995) procedure.

Table A.6: Impact of 10 p.p. increase in GP practice propensity to refer to NPMH on various combinations of mental health care in the first 12, 24, 48 and 96 months after the GP appointment

	12 months	24 months	48 months	96 months
NPMH, Basic and Specialised MH	0.05%*	0.10%*	0.11%*	-0.01%
	(0.01)	(0.01)	(0.02)	(0.02)
NPMH and Basic MH	0.33%**	0.44%**	0.44%**	0.29%**
	(0.02)	(0.02)	(0.03)	(0.03)
NPMH and Specialised MH	0.37%**	0.51%**	0.43%**	0.27%**
	(0.02)	(0.02)	(0.03)	(0.03)
NPMH only	3.76%**	3.89%**	3.36%**	2.92%**
	(0.05)	(0.05)	(0.06)	(0.06)
Basic and Specialised MH	-0.07%*	-0.13%*	-0.14%*	-0.09%*
	(0.01)	(0.02)	(0.02)	(0.02)
Basic MH only	-0.67%**	-0.77%**	-0.68%**	-0.45%**
	(0.04)	(0.04)	(0.04)	(0.04)
Specialised MH only	-0.49%**	-0.57%**	-0.45%**	-0.27%**
	(0.04)	(0.04)	(0.04)	(0.04)

Standard errors shown in parentheses; *significant at a 5% significance level; **significant at a 1% significance level, significance level reported after correcting for multiple hypothesis testing using the Benjamini & Hochberg (1995) procedure.

Table A.7: Robustness of impact of 10 p.p. increase in GP practice propensity to refer to NPMH on health care utilization in 12 months before and 12 months after initial GP appointment

	Baseline	Without GP FE	Exclude zero propensity	1 month propensity	Pre-2017 sample
			12 pre-months		
P(Basic MH)	-0.04%	-0.07%	-0.03%	-0.03%	-0.01%
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)
P(Specialized MH)	0.02%	0.04%	-0.02%	-0.01%	0.03%
	(0.03)	(0.03)	(0.03)	(0.02)	(0.04)
P(Pharmaceuticals)	0.06%	-0.01%	0.12%*	0.04%	0.09%
	(0.05)	(0.06)	(0.06)	(0.03)	(0.06)
			12 post months		
P(NPMH visit)	5.01%**	7.71%**	3.19%**	2.72%**	5.09%**
	(0.05)	(0.11)	(0.07)	(0.04)	(0.07)
P(GP visit>1)	0.18%*	0.00%	0.28%**	0.08%	0.43%**
	(0.08)	(0.17)	(0.09)	(0.06)	(0.10)
P(Basic MH)	-0.33%**	-0.41%**	-0.26%**	-0.16%**	-0.42%**
	(0.04)	(0.09)	(0.05)	(0.03)	(0.05)
P(Specialized MH)	-0.10%	-0.03%	-0.05%	-0.07%*	-0.11%
	(0.04)	(0.06)	(0.05)	(0.03)	(0.06)
P(Pharmaceuticals)	-0.03%	-0.22%*	0.00%	-0.05%	-0.01%
	(0.06)	(0.10)	(0.07)	(0.04)	(0.08)

Standard errors shown in parentheses; *significant at a 5% significance level; **significant at a 1% significance level, significance level reported after correcting for multiple hypothesis testing using the Benjamini & Hochberg (1995) procedure.

Table A.8: Impact of 10 p.p. increase in GP practice propensity to refer to NPMH on quantity of health care in the first 12, 24, 48 and 96 months after the GP appointment

	12 pre-months	12 months	24 months	48 months	96 months
Num NPMH visits		0.22**	0.24**	0.24**	0.18**
		(0.00)	(0.00)	(0.01)	(0.01)
Num GP visit		-0.01	-0.00	0.01	0.02
		(0.01)	(0.01)	(0.01)	(0.01)
Basic MH spending	-0.02	-2.14**	-2.39**	-2.20*	-2.23*
	(0.20)	(0.52)	(0.62)	(0.81)	(0.96)
Specialized MH mins	-0.78	-0.35	0.42	-1.58	-1.81
	(0.60)	(0.78)	(1.32)	(2.10)	(2.92)
Num Pharmaceuticals	0.04	0.02	0.06	0.14	0.29
	(0.02)	(0.02)	(0.04)	(0.08)	(0.13)
Physical spending	-3.24	13.49	19.62	27.97	51.18
	(8.78)	(11.13)	(16.04)	(25.39)	(36.29)

Standard errors shown in parentheses; *significant at a 5% significance level; **significant at a 1% significance level, significance level reported after correcting for multiple hypothesis testing using the Benjamini & Hochberg (1995) procedure.

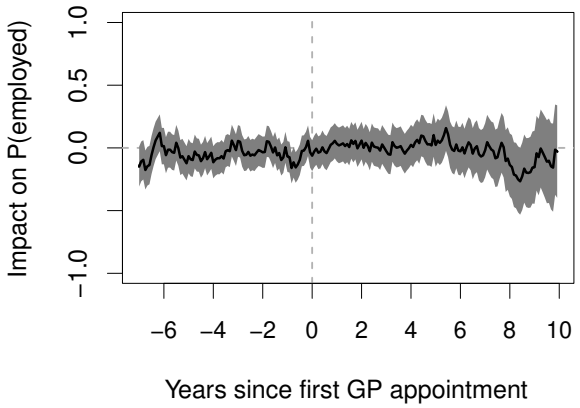
Table A.9: Reduced form impact of 10 p.p. increase in GP propensity to refer to talk therapy on self-assessed and physician-assessed health in the 48 months prior to the GP appointment and 48 months after the initial GP appointment

	48 pre-months	48 post-months
Self-assessed health[a]:		
General health	-0.01	0.06
	(0.10)	(0.11)
Mental health	-0.05	0.03
	(0.06)	(0.08)
Risk of anxiety disorder or depression	-0.12	-0.14
	(0.10)	(0.12)
Loneliness	-0.14	0.05
	(0.10)	(0.12)
Sample size	24,394	19,301
Physician-assessed health[b]:		
At start of treatment	-0.07	-0.08
	(0.07)	(0.07)
At end of treatment	-0.02	-0.03
	(0.07)	(0.07)
Highest assessed	0.06	-0.01
	(0.07)	(0.07)
Sample size	95,865	91,045

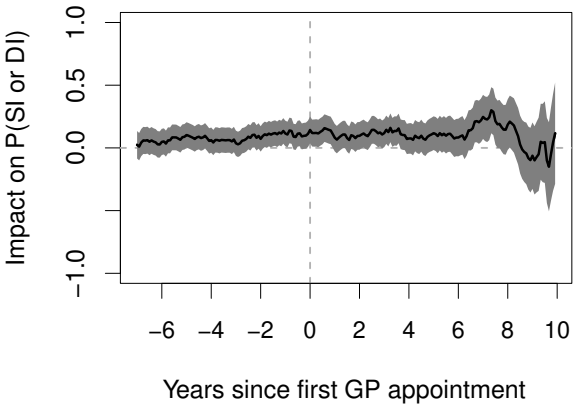
Standard errors shown in parentheses; *significant at 5%; **significant at 1%; significance levels reported after correction for multiple hypothesis testing with the Benjamini & Hochberg (1995) procedure.

Figure A.1: IV: Reduced form impact of a 10 p.p increase in GP practice propensity to refer to an NPMH on employment, DI/SI benefit receipt and earnings measured starting 7 years before the initial GP appointment until 10 years after the initial GP appointment.

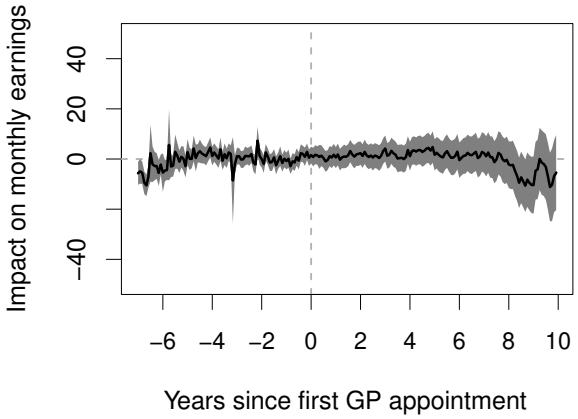
(a) Employment



(b) DI/SI benefit receipt



(c) Earnings



A.3 Alternative estimation approach: First hire of an NPMH as Regression-Discontinuity-in-Time

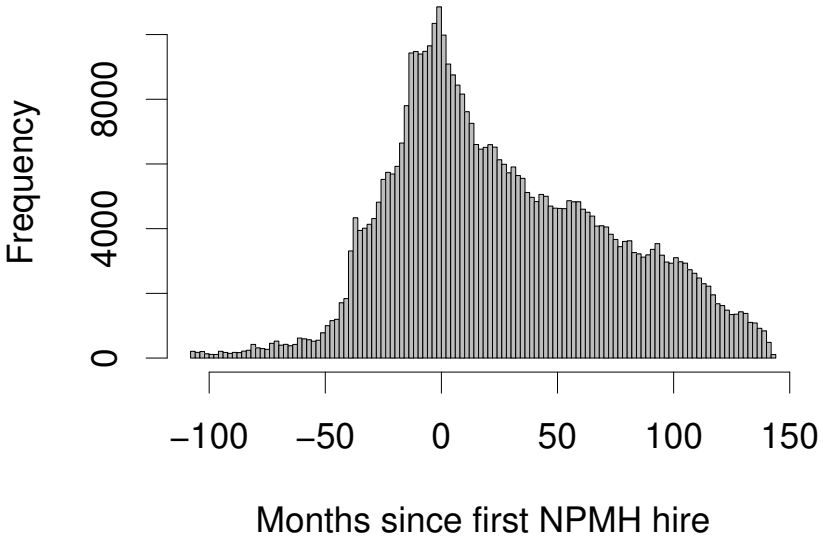
The main IV analyses exploits variation in GP propensity to refer to an NPMH to estimate the impact of talk therapy on health care utilization and labor market outcomes. The exclusion assumption imposes that differences in GP propensity only affect patients through the probability to receive talk therapy from an NPMH. This section loosens this assumption by only comparing patients right before and right after an NPMH was first hired in a Regression-Discontinuity-in-Time framework.

As a response to increased funding for NPMHs, an increasing number of GP practices have hired NPMHs. I reconstruct the moment a GP practice first hired an NPMH by using the NPMH utilization rates at the GP practice level calculated for the IV strategy and determining the first month in which patients at the practice see an NPMH.

To illustrate, Figure 5 in the main text showed the utilization rates at 4 different practices. For the practice in panels (a) and (b), a utilization of zero is observed in 2011 and 2012, after which the utilization rate increases to some positive amount. For both of these GP practices, the first moment an NPMH is hired can thus be determined to be January 2013. For the GP practice in panel (c), the utilization rate is positive throughout the entire 2011-2024 window, and hence, the moment an NPMH is first hired at this practice cannot be determined. Similarly, the GP practice in panel (d) never refers to an NPMH, and hence I again do not observe the moment an NPMH was first hired. Using the GP practice level utilization rates, I determine the first moment of hire for all 609 Nivel-affiliated GP practices. For 269 of these practices, the first hiring of an NPMH is observed.

Given the months of the first hire, I estimate the impact of talk therapy provided by NPMHs by comparing individuals who contact their GP right before or right after this moment. The intuition behind this is that individuals contacting their GP right after have a higher probability of receiving treatment from an NPMH while being similar in all other characteristics. One potential issue could be that hiring an NPMH directly impacts the inflow of patients. Patients could, for example, contact their GP after hearing about

Figure A.2: Distribution of patients contacting their GP relative to the month an NPMH is first hired



this treatment option, or GPs could proactively contact individuals they deem suitable for treatment by NPMHs. To determine whether this is the case, Figure A.2 shows the distribution of individuals contacting their GP relative to the month an NPMH is first hired.

As seen in Figure A.2, there is no increase in the number of patients contacting their GP in the months after an NPMH was first hired. Note that the shape of the distribution, with more mass around zero and a long right tail, is caused by the time window for which the Nivel data is available and the time at which the NPMHs are first hired. The Nivel data is available between 2011 and 2024, and most hirings occur between 2012 and 2017.

To further test whether there are any discontinuities around the threshold, Table A.10 additionally shows that there are no differences in background characteristics of patients contacting their GP right before and right after the NPMH is hired. These results thus confirm that individuals seeking care right before and right after the NPMH is first hired, are likely to be similar in observable and unobservable characteristics but potentially face different odds of receiving treatment from an NPMH.

I exploit the discontinuity in treatment availability through a RDiT design approach. In this approach, the running variable is the month at which an individual contacts their

Table A.10: Balance of individual characteristics around threshold

Characteristic	Estimated discontinuity
Age	-0.33
	(0.23)
Male	-0.24%
	(0.49)
Native	0.44%
	(0.40)
Education unknown	-0.27%
	(0.49)
Education low	0.53%
	(0.60)
Education middle	0.08%
	(0.63)
Education high	-0.61%
	(0.57)

Standard errors shown in parentheses; *significant at a 5% significance level; **significant at a 1% significance level, significance level reported after correcting for multiple hypothesis testing using the Benjamini & Hochberg (1995) procedure.

GP, relative to the month at which the NPHM is hired. This method thus compares individuals contacting their GP right before the NPMH is hired to those contacting their GP right after, while flexibly controlling for any time effects. The baseline specification looks as follows:

$$H_{it} = \alpha + f(t) + I(t > 0)g(t) + \delta I(t \geq 0) + \varepsilon_{it} \tag{A.1}$$

with H_{it} the outcome of interest for individual i who contacts their GP t months after the NPMH is hired. $f(t)$ and $g(t)$ flexibly control for the impact of time relative to the month of hire, and $I(t > 0$ indicates whether an individual has contacted their GP after the NPMH is hired. δ is thus the parameter of interest, capturing a jump in the outcome of interest at the month the NPMH is hired. In the baseline specifications, $f(t)$ and $g(t)$ are quadratic functions, and a triangular kernel is used around the threshold. The Imbens Kalyanaraman optimal bandwidth is used in all regressions.

H_{it} is measured over a certain time period, ranging from the first 12 months after the initial GP appointment to the first 96 months after the initial GP appointment. Similar

to the IV estimation, I also consider outcomes measured in the 12 months preceding the initial GP appointment. As patients have not seen their GP yet, there should be no impact of the NPMH and these estimates thus serve as placebo test.

Results

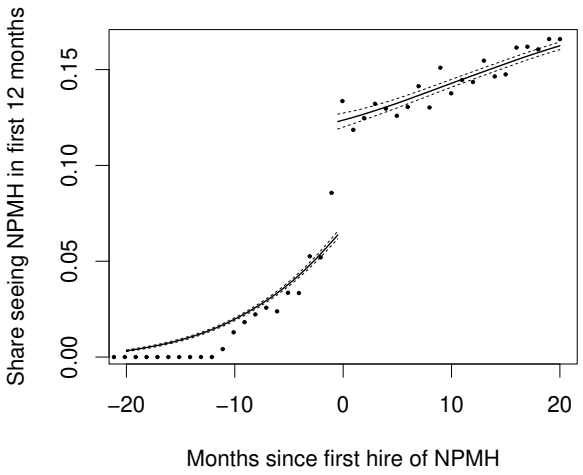
Figure A.3 shows the RDiT estimate of the impact of the first hiring of an NPMH on health care utilization in the first 12 months after the initial GP appointment. Panel (a) shows a significant increase in the probability of receiving talk therapy from an NPMH of approximately 5 p.p. and a small decrease in the probability of seeing the GP. Basic mental health care, specialised mental health care and the use of pharmaceuticals is not affected by the initial hiring of an NPMH.

To assess the dynamics of the impact of the introduction of the NPMH on all possible mental health care treatment options, Table A.11 shows the RD estimates when using NPMH visits, GP visits, basic mental health care, specialised mental health care, or pharmaceuticals as outcomes of interest. Similar to the IV estimation, the various columns use different time windows relative to the GP appointment. Column (1) of the table is a placebo test, as it examines differences in treatment uptake in the 12 months before the GP appointment. By construction, the individuals in the sample have no GP or NPMH visits in this period. Reassuringly, the RDiT analyses also show no significant differences in the uptake of any of the other treatment options.

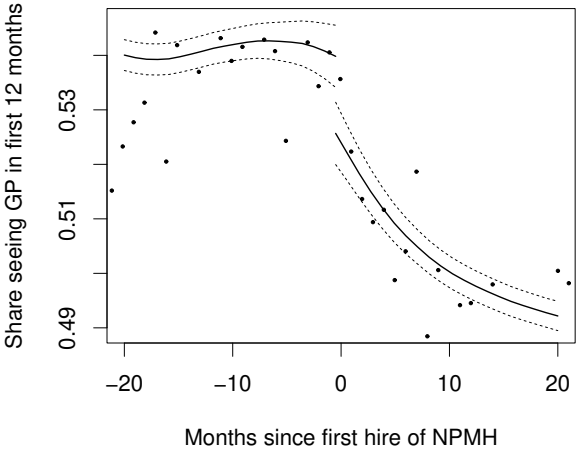
Columns (2) to (5) show differences in treatment uptake in the first 12, 24, 48 and 96 months after the initial GP appointment. For all time windows, there is a significant increase in the probability of receiving treatment by an NPMH. The magnitude of this increase decreases from 6 percentage points to 5 percentage points as the observation window becomes longer. This implies that some individuals contacting their GP before the first hire of the NPMH will eventually receive treatment from an NPMH. They might see their GP again a couple of months after the initial GP appointment (after the NPMH is hired) and be referred to the NPMH. However, for the majority of the individuals contacting their GP before the NPMH is hired, catch-up of treatment does not occur.

Figure A.3: RDiT estimate of first hiring of NPMH on health care utilization in the first 12 months after the initial GP appointment

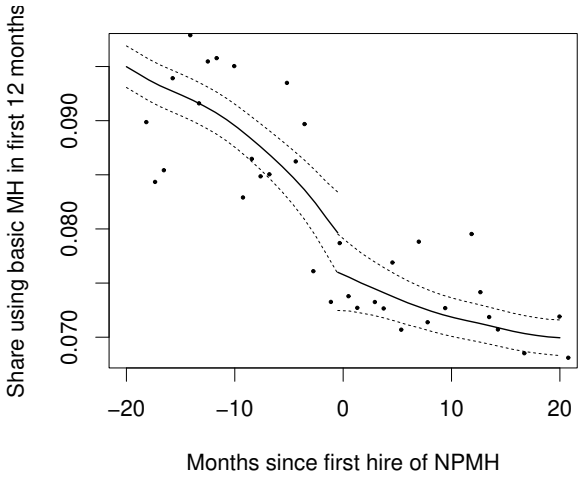
(a) NPMH



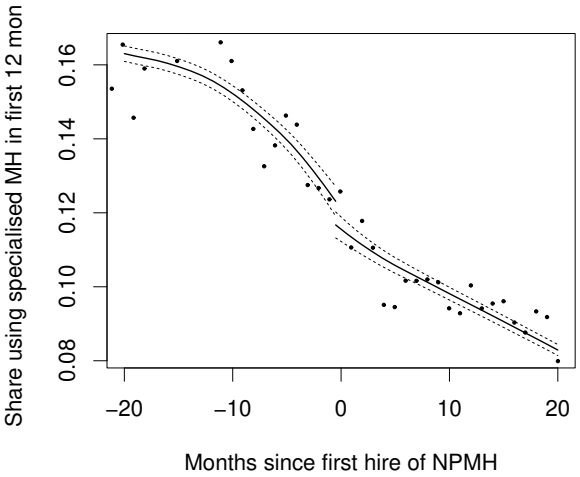
(b) GP



(c) Basic mental health care



(d) Specialised mental health care



(e) Pharmaceuticals

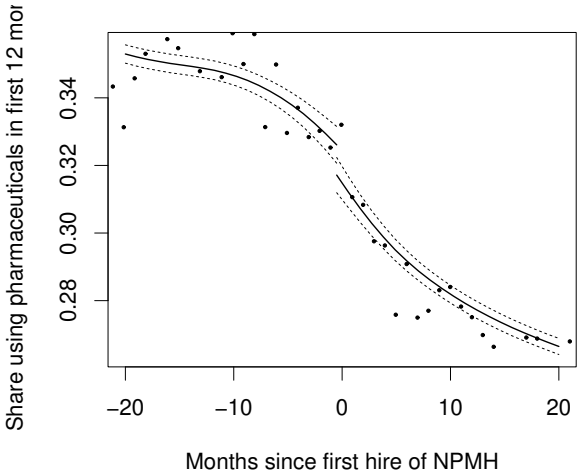


Table A.11 furthermore shows that individuals contacting their GP after the NPMH is hired are less likely to see their GP in later months and less likely to receive treatment in

Table A.11: RDiT estimates of differences in treatment uptake in 12 months prior to GP appointment and the first 12, 24, 48 and 96 months after the GP appointment

	12 pre-months	12 months	24 months	48 months	96 months
P(GP visit>1)		-1.41%*	-1.54%*	-1.31%*	-1.14%*
		(0.50)	(0.49)	(0.46)	(0.44)
P(NPMH visit)		5.92%**	5.53%**	5.39%**	5.43%**
		(0.29)	(0.32)	(0.37)	(0.41)
P(Basic MH)	-0.33%	-0.37%	-0.47%*	-0.30%*	-0.06%
	(0.16)	(0.28)	(0.18)	(0.13)	(0.16)
P(Specialized MH)	-0.09%	-0.64%	-0.53%	-0.39%	-0.48%
	(0.25)	(0.33)	(0.36)	(0.38)	(0.40)
P(Pharmaceuticals)	-0.83%	-0.89%	-0.86%	-0.88%	-0.74%
	(0.42)	(0.47)	(0.49)	(0.50)	(0.50)
Num GP visit		-0.10**	-0.15**	-0.07	-0.04
		(0.03)	(0.04)	(0.06)	(0.09)
Num NPMH visits		0.25**	0.24**	0.27**	0.33**
		(0.02)	(0.02)	(0.06)	(0.09)
Basic MH spending	-2.09	-2.55	-4.61*	-2.24	0.81
	(0.96)	(2.88)	(1.98)	(1.50)	(1.67)
Specialized MH mins	-0.92	-2.16	3.16	9.97	27.86
	(4.52)	(5.45)	(9.50)	(15.27)	(22.63)
Num Pharmaceuticals	-0.20	-0.14	-0.12	0.01	0.16
	(0.15)	(0.16)	(0.31)	(0.61)	(1.01)
Physical spending	13.90	9.55	-61.42	12.12	15.87
	(60.49)	(69.26)	(50.76)	(37.74)	(46.06)

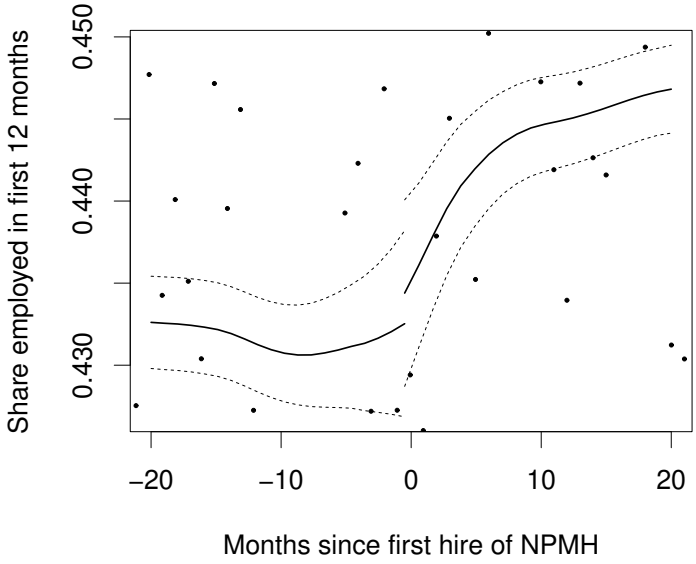
Standard errors shown in parentheses; *significant at a 5% significance level; **significant at a 1% significance level, significance level reported after correcting for multiple hypothesis testing using the Benjamini & Hochberg (1995) procedure.

basic mental health care. The magnitude of these changes to existing treatment options, however, is much smaller than the increased take-up of treatment by NPMH. This implies that for most patients, talk therapy provided by NPMHs acts as a complement to the already existing treatment options. The impact on the extensive margin of health care receipt are mirrored by changes in the intensive margin, as shown in rows 6 to 11.

To determine whether the increased uptake of talk therapy provided by NPMHs also translates into improved labor market outcomes for patients, Figure A.4 shows RD estimates using employment in the first 12 months after the initial GP appointment as the outcome of interest. Note that these estimates correspond to a sharp RD setup; they do not show the impact of receiving talk therapy, but rather the impact of contacting a GP right after the initial hiring of the NPMH. As can be seen from the figure, the

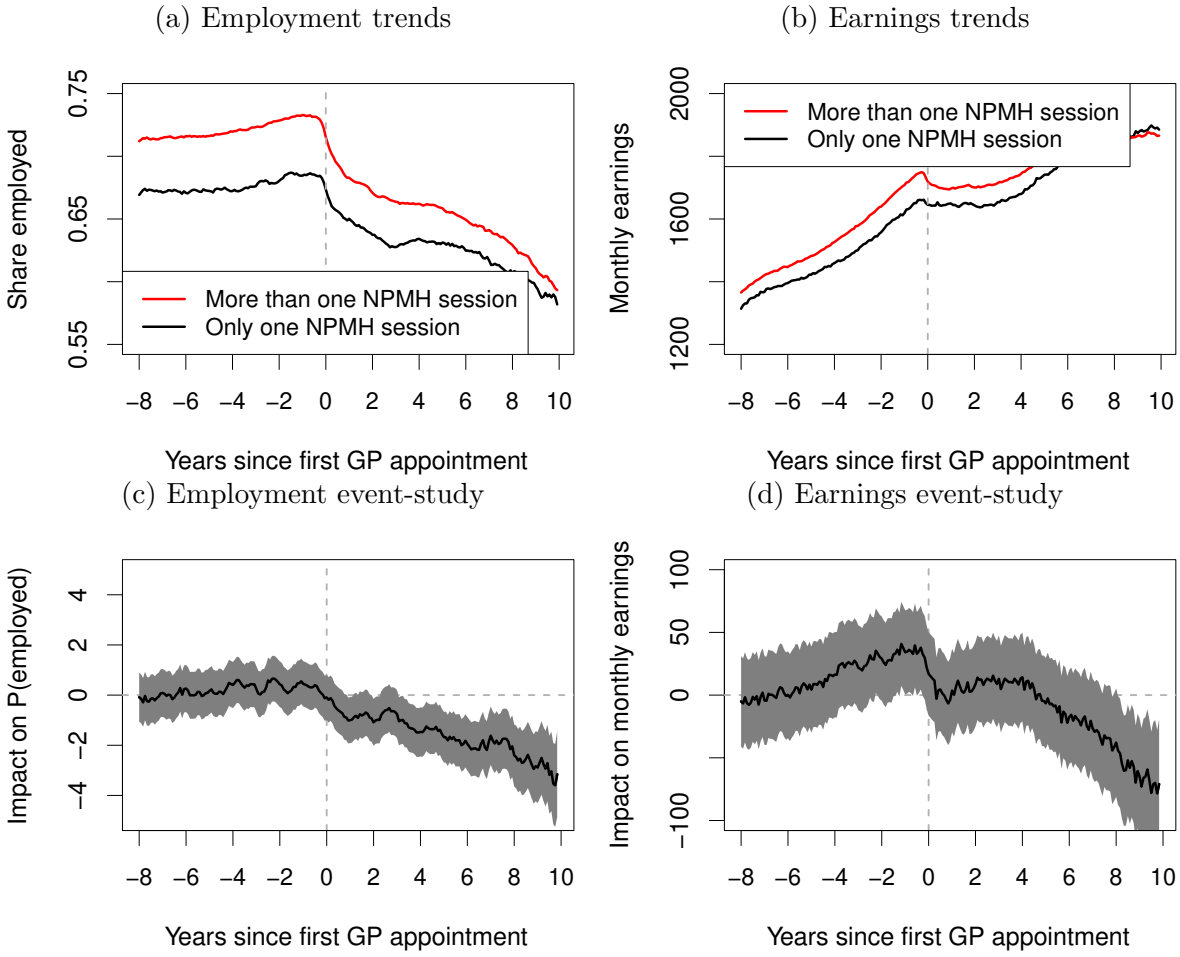
estimates are insignificantly different from zero. The limited sample used in this analysis, paired with relatively high variance in employment, does, however, render the RD setup underpowered to detect sizeable employment impacts.

Figure A.4: RDiT estimate on the share of patients employed in the first 12 months after the initial GP appointment



A.4 Replication of Rzepnicka et al. (2025)

Figure A.5: Replication of Rzepnicka et al. (2025) showing trends in labor market outcomes and corresponding event-study estimates comparing individuals that stop treatment after one session with those that continue receiving treatment



B Inequality in treatment uptake

Previous research has shown that the uptake of treatment for (mental) health problems can differ significantly based on individual characteristics such as migration background and education. Sentell et al. (2007); Cook et al. (2017) for example find that non-white individuals in the US are up to 80% less likely to use mental health care, compared to their white counterparts. Talk therapy being offered by NPMHs was partly introduced to decrease these treatment gaps by increasing uptake of treatment among groups that were being undertreated. To assess whether it has been effective in doing so, this section will first estimate these disparities in mental health care utilization.

The descriptive statistics in Table 1 potentially hint at moderate differences in treatment uptake in the Netherlands based on gender and education level. However, to accurately measure disparities in treatment uptake, both the utilization of care and the need for care should be included. To illustrate, the descriptive statistics show that individuals with a migration background are likelier to use mental health care than those without a migration background. However, this does not necessarily imply that individuals without a migration background use disproportionally too little mental health care. It could be the case that individuals with a migration background have a higher (or lower) need for care and should thus have a higher (or lower) utilization of care.

To determine whether disparities in mental health care uptake exist in the Netherlands, I analyze the annual health care expenditure and survey measures on health for the full population in 2012 (column (1) of Table 1) and estimate the following equation:

$$H_i = \alpha + \beta X_i + \delta Z_i + \varepsilon \tag{B.1}$$

with H_i an indicator for using either basic or specialized mental health care[17] X_i indicators for age, gender, migration background and education level and Z_i survey measures of health. To determine the influence of differences in health, Equation B.1 is estimated while including and excluding the survey health measures Z_i and while conditioning on the subsamples with a high risk of either psychological disorder, anxiety and depression, or loneliness.

Table B.1 shows the estimated differences in mental health care utilization. Column (1) shows that when not controlling for differences in underlying (self-reported) health, older individuals, males, Dutch natives, and higher-educated individuals are less likely to use mental health care. The estimated differences are sizeable, ranging from 0.51 to 2.09 percentage points, or 12 to 50% of the baseline utilization rate.

Column (2) shows that for some background characteristics, differences in self-reported health partly explain the differences in health care utilization, while other differences in self-reported health mask the gap in utilization. More specifically, the substantial differ-

[17]Results using only basic mental health care or only specialized mental health care are similar.

Table B.1: Differences in mental health care utilization rates in 2012 conditional and unconditional on self-reported mental health

	Unconditional regression	Conditional on self-reported mental health
Aged $< 45^a$	1.84%**	1.73%**
	(0.10)	(0.09)
Aged $> 65^a$	-2.09%**	-2.33%**
	(0.09)	(0.09)
Male^b	-1.36%**	-0.58%**
	(0.07)	(0.07)
Dutch native^c	-0.61%**	0.92%**
	(0.10)	(0.10)
Low education^d	0.59%**	-1.37%**
	(0.15)	(0.15)
Middle education^d	0.51%**	-0.44%**
	(0.12)	(0.12)
Average utilization rate	4.4%	4.4%
Number of individuals	365,281	350,791

Standard errors shown in parentheses; *significant at a 10% significance level; **significant at a 5% significance level; (a) Omitted age category is individuals aged between 45 and 65. Note that individuals 18 or younger are not included in survey data; (b) Omitted gender is female; (c) Omitted migration background is first or second-generation migrant; (d) Omitted educational level is high educated.

ence in utilization between age groups is largely unaffected by differences in self-reported health. For both regressions, individuals aged 18-45 are approximately two p.p. (30%) likelier to use care than those aged 45-65, while those older than 65 are two p.p. (30%) less likely to use mental health care compared to those aged 45-65. Higher utilization rates at younger ages align with literature showing that the onset of many mental health issues tends to be during early adulthood (OECD, 2014; Jones, 2013). However, the gaps in treatment conditional on self-reported health could indicate undertreatment among older cohorts. Alternatively, these individuals might have already been treated at a younger age.

When looking at gender, differences in self-reported health explain approximately half of the difference in utilization. When accounting for self-reported health, males are roughly 10% less likely to use mental health care compared to females. These differences

largely align with literature showing lower treatment uptake rates for men (Hauenstein et al., 2006).

For differences in mental health care utilization based on migration background and educational attainment, differences in self-reported health have a significant impact and flip the direction of the disparity. Without controlling for self-reported health, Dutch natives and higher educated individuals are 0.6 p.p (15%) less likely to use mental health care than their counterparts. However, these groups tend to use relatively more health care when controlling for self-reported health. More specifically, Dutch natives are 0.9 p.p. (20%) likelier to use mental health care than non-natives, whereas higher educated individuals are 1.4 p.p. (30%) likelier to use care. This impact of differences in health care need is confirmed by regressions on subsamples with elevated risk of mental health issues, reported in Appendix Table B.2.

It should be noted that the survey measures of health could be impacted by the utilization of care, in which case they would be bad controls. However, the analysis does show that when examining differences in health care utilization, controlling for self-reported health measures can have large implications. While treatment disparities are less substantial than those reported for the US, differences in treatment uptake for mental health problems are still significant. Older individuals, males, individuals with a migration background, and those with lower educational attainment are between 10 and 30% less likely to use care.[18] Given these significant disparities, it is essential to determine whether an expansion of treatment options closes part of the treatment gap or only further widens the gap. In the next Section, I will therefore first estimate the impact of talk therapy on the full population, before turning to heterogeneity in the impact by age, gender, education level, and migration background.

[18]For age, the difference can be up to 130% when comparing those below the age of 45 to those above the age of 65. However, this age difference might be explained by the fact that older individuals had already received treatment when they were younger.

Table B.2: Differences in mental health care utilization rates in 2012 conditional and unconditional on self-reported mental health

	Unconditional regression	Conditional on self-reported mental health	High risk psychological disorder	High risk anxiety and depression	High risk loneliness
Aged $< 45^a$	1.84%**	1.73%**	7.35%**	3.73%**	6.17%**
	(0.10)	(0.09)	(0.86)	(0.21)	(0.55)
Aged $> 65^a$	-2.09%**	-2.33%**	-11.14%**	-5.61%**	-8.02%**
	(0.09)	(0.09)	(0.87)	(0.19)	(0.50)
Male^b	-1.36%**	-0.58%**	-0.45%	-0.98%**	-1.35%**
	(0.07)	(0.07)	(0.67)	(0.15)	(0.39)
Dutch native^c	-0.61%**	0.92%**	7.89%**	1.32%**	4.20%**
	(0.10)	(0.10)	(0.74)	(0.20)	(0.47)
Low education^d	0.59%**	-1.37%**	-8.56%**	-1.53%**	-2.20%**
	(0.15)	(0.15)	(1.27)	(0.31)	(0.81)
Middle education^d	0.51%**	-0.44%**	-3.68%**	0.03%	-1.16%
	(0.12)	(0.12)	(1.18)	(0.26)	(0.74)
Average utilization rate	4.4%	4.4%	23.6%	8.5%	12.0%
Number of individuals	365,281	350,791	17,048	139,495	26,853

Standard errors shown in parentheses; *significant at a 10% significance level; **significant at a 5% significance level; (a) Omitted age category is individuals aged between 45 and 65. Note that individuals 18 or younger are not included in survey data; (b) Omitted gender is female; (c) Omitted migration background is first or second-generation migrant; (d) Omitted educational level is high educated.