Intergenerational Effects of Sick Leave on Child Human Capital

Julie Riise, Barton Willage, and Alexander Willen*

July 2024

Abstract

This paper examines the intergenerational impact of one of the largest employment protection and income replacement programs in the world: sick leave. To do so, we exploit random assignment of patients to doctors in the Norwegian health care system and use the fact that some physicians will be more lenient than others in terms of issuing sick leave certificates. Using detailed administrative data that enables us to match patients to their physicians as well as to their children, we show that random assignment to a physician who is more lenient in issuing sick leave certificates negatively impacts the human capital development of the patients' children across the educational system; from GPA in early adolescence to enrollment in higher education. In terms of mechanisms, we find that sick leave has negative long-term effects on parental earnings, makes parents increasingly dependent on the social insurance system, and negatively affects their mental health.

JEL Codes: I10, I20, J20, J60
Keywords: Sick Leave, Human Capital Development, Intergenerational Links

* Willén: Department of Economics, Norwegian School of Economics, CESifo, UCLS, and IZA (alexander.willen@nhh.no). Willage: Department of Economics, University of Deleware and NBER (willage@udel.edu). Riise: Department of Economics, University of Bergen, and IZA (julie.riise@uib.no). Willén gratefully acknowledges financial support from the Research Council of Norway through its Centers of Excellence Scheme, FAIR project no. 262675.

1 Introduction

Children are highly affected by their home environments, and it is well established that the structure and resources of families influence the human capital development of children (e.g., Knudsen et al. (2006); Cunha et al. (2006); Heckman and Mosso (2014); Carneiro et al. (2022)). This has motivated a strong emphasis in the academic and policy literatures on the value of a social safety net that protects children from abrupt changes and transitory shocks during their childhood. However, the direction of the intergenerational effect of parental welfare participation on child development is theoretically ambiguous. On the one hand, basic labor market support systems such as sick leave schemes and unemployment insurance programs may benefit children by providing financial security, reducing stress, and improving the home environment. On the other hand, if higher participation induces increased welfare dependence, negatively impacts the parents' career trajectory, and transmits adverse role model/perception signals to the child, this could lead to worse child outcomes and increased government expenditures.

Disentangling the intergenerational impact of employment protection and income replacement programs on child development is challenging. Not only does it require detailed multi-generational data that spans many years, but it also requires variation in parental use of the social safety net that is uncorrelated with other determinants of child development (e.g., the underlying shock that leads to use of the social safety net). In this paper, we overcome these challenges and provide the first quasi-experimental evidence on the overall intergenerational impact of one of the largest employment protection and income replacement programs in the world: sick leave. This scheme is considered a core feature of most labor market systems and accounts for a non-negligible share of national GDP spending across the OECD.[1]

To perform our analysis, we take advantage of two unique features of the Norwegian sick leave system. First, sick leave certificates must be issued by licensed health personnel, and General Practitioners (GPs) certify the overwhelming majority of sick leave days. Second, when GPs retire, move, reduce their practice, or exit the market, existing patients must be randomly reallocated to new physicians.[2] Because GPs differ in their leniency to provide sick leave certificates to their patients, we can construct a measure of GP leniency and exploit this randomization process to overcome any endogeneity concerns and obtain quasi-experimental variation in parental sick leave. The randomization of patients to new GPs

[1] Ranging from a low of 0.2 percent of GDP in Greece to more than 2 percent in countries such as Germany, Norway, and the Netherlands; see the Organisation for Economic Co-operation and Development Social Expenditure Database (https://www.oecd.org/social/expenditure).

[2] This applies to cases when not all patients are transferred to the same new GP.

is performed by the Norwegian Health Economics Administration through an automated computer system and is regulated by Norwegian law (FOR-2012-08-29-842, Section 35).[3] This ensures exogenous variation in GP-patient match in our setting.

In all our specifications, we include previous GP fixed effects and control for baseline sick leave usage. The thought experiment underlying our empirical approach is, therefore, to compare the human capital development of children whose parents initially had the same GP, lived in the same area, and had the same baseline sick leave usage, but then were randomly reallocated to new, and different, GPs due to external factors outside their control. The legally mandated random assignment mechanism breaks any correlation between sick leave and unobserved determinants of individual outcomes, and enables us to isolate the intergenerational impact of sick leave leniency on child human capital development holding all other factors constant (such as, for example, individual health status).

After identifying the sick leave leniency of GPs and examining the impact on child development, we study a range of mechanisms that may help explain the intergenerational effects that we uncover. We do this by linking the patient data to a rich set of labor market, welfare participation, and mental health outcomes. This allows us to identify the impact of GP sick leave leniency on the parents themselves – and through which of these channels the effects on children may operate. To obtain a comprehensive understanding of the impact of sick leave leniency on patient outcomes, we estimate effects both in the immediate short-run and in the slightly longer-run (five years after assignment).

Our leniency measure is identified under the assumption that there is no systematic sorting of patients to GPs as a function of GP sick leave leniency. In other words, patients who experienced the same GP exit or list reduction cannot be systematically sorted across new GPs based on patient characteristics that also are correlated with their outcomes. In theory, the validity of this assumption follows directly from the legally-mandated computerized randomization procedure used to reallocate patients across GPs. In practice, it is possible to obtain suggestive evidence on the validity of this assumption by examining if the leniency of the exogenously-assigned GP is uncorrelated with observed patient characteristics. Using a rich set of patient characteristics, *including variables related to prior health care utilization, labor market outcomes, spousal characteristics, and baseline demographics*, we find strong evidence in support of this assumption. We also conduct a rich set of placebo tests, robustness checks, and sensitivity analyses, to rule out potential confounders and other threats to identification; we discuss all of the results from these exercises in the robustness section after

[3]The law does not specify the method used for randomization, and information about the computer system used for randomization was obtained through private correspondence with the Norwegian Health Economics Association.

our main results.

A final and very important concern is whether a GP's leniency in issuing sick leave certificates is linked to other aspects of their practice style or quality. While this would not undermine the causal identification in our study, it would complicate the interpretation of our results. Specifically, it would suggest that part of the observed effect might be driven by unmeasured GP characteristics correlated with sick leave leniency that independently influence the human capital development of the patients' children. We take this concern very seriously and conduct a thorough set of analyses to examine any potential links between sick leave leniency and GP practice style or quality. The results are clear: we find no evidence of any relationship between sick leave leniency and key indicators of GP practice style or quality, including measures like GP value-added, treatment intensity (as proxied by reimbursements per visit), and the number of patients a GP has. This indicates that lenient sick leave certification is not associated with a GP's ability to improve patient health. This is consistent with prior Norwegian studies that also have found no link between leniency and GP quality (e.g., Markussen (2012); Markussen and Røed (2017)).

Furthermore, we find no evidence that being assigned to a lenient GP affects patient healthcare usage. We examine a wide range of health care outcomes, including short- and long-term inpatient visits, ER visits, and the likelihood of GP check-ups. Across these measures, there is no correlation with GP leniency. These findings align with previous research in Norway (e.g., Riise et al. (2022); Ginja et al. (2024)) and indicate that sick leave decisions are likely unrelated to both patient healthcare utilization and GP practice style. This is expected, given the subjective nature of sick leave certification, especially for conditions that drive variation in leniency—a point discussed in depth later.

While it is theoretically possible that an unobserved confounder exists—one that correlates with GP leniency but not with GP practice style, quality, or patient healthcare use, and selectively influences the human capital development of patients' children—we find no evidence in support of this. Such a variable would also need to disproportionately affect patients with less severe health conditions that require subjective interpretation of sick leave (conditions that drive our variation in leniency, musculoskeletal conditions and psychological conditions). We find no plausible candidate for such a confounder, and it would also run counter to the evidence presented in the prior literature. Additional placebo tests, robustness checks, and heterogeneity analysis provide further evidence in support of our causal interpretation. Thus, we interpret our results as operating through the sick leave leniency channel.

We present three key results. First, we show that there is considerable variation in sick leave as a function of GP leniency. Being assigned to a GP who is located 1 SD above the mean leniency generates an *additional* 4.5 weeks (or 22 days) of paid sick leave relative to being assigned to a GP who is located 1 SD below the mean leniency on top of what is already taken. This is a substantial amount, equivalent to more than a month of full-time employment, or 10 percent of a normal work year in the country (or a 25 percent difference relative to the mean). Variation in leniency is much more pronounced for hard-to-verify musculoskeletal and psychological causes for which GPs arguably have more freedom in their sick leave decisions, both at the intensive as well as the extensive margin.

Second, we find economically and statistically significant negative effects of parental sick leave across the child's human capital development. Specifically, a one standard deviation increase in the leniency of a parent's GP is associated with a decline in compulsory school GPA and high school GPA by 1.6 percent of a standard deviation. This is an economically meaningful decline, though smaller than the effects of targeted education inventions such as class size reductions or teacher quality improvements (e.g., Fredriksson et al. (2013); Chetty et al. (2014a); Krueger and Whitmore (2001)), in line with modest changes in household resources (e.g., Dahl and Lochner (2012)), and much smaller than severe parental shocks on child human capital development such as job loss (e.g., Carneiro et al. (2022)).

Third, we find that the effects on the children's human capital development are not only operating on the intensive margin of educational attainment, but also are present on the extensive margin. Specifically, in response to the parent being assigned a GP who is 1 SD more lenient, we find a reduction in the probability of graduating high school and of attending college by approximately 0.8 percentage points. Thus, the impact of parental employment protection on child development does not only affect the quality of the human capital that children accumulates, but also the quantity.

In terms of mechanisms, we find that assignment to a more lenient GP leads parents to earn lower wages, become more dependent on the social welfare system, and receive more in total welfare benefits – both in the short- and the long-run. We show that the sum of these effects generate a decline in the mental health of parents in the longer-run. Thus, assignment to a more lenient GP does not only cause parents to stay at home for extended sick leave spells, but also impacts other dimensions of their work-life for up to five years after assignment. We hypothesize that the aggregation of all these labor market, mental health, and welfare dependence effects help explain part of the effect on the human capital development of children.

Even though the child effects we identify operate through more channels than just parental income, it is instructive to think about how our intergenerational effects compare to other studies that have looked at income changes. This provides us with bounds to consider the credibility of the magnitude of the effects that we identify. Using non-linear changes in the EITC, Dahl and Lochner (2012) finds that a 4 percent increase in parental income generated a 6 percent of a standard deviation change in child educational performance. If we scale our child effects by the impact on parental income, we derive slightly larger but very similar effects: a 1 percent change in parental income (or 0.8 percent if we account for the increase in welfare benefit payout that we identify) generates a 1.6 percent of a standard deviation change in student GPA. That our effects are slightly larger is consistent with the idea that our results operate through more channels than just family income (such as, for example, mental health and welfare dependence).

The main contribution of our paper is to exploit exogenous variation in GP sick leave leniency and combine this with detailed register data on patients as well as their children to provide novel evidence on how parental take-up of one of the largest employment protection and income replacement programs in the world impacts the human capital development of children across their childhood. This allows us to advance the existing literature in several distinct ways.

First, there is an impressive literature on the life-cycle approach to skill formation, focusing on the interaction between parental investments and childhood development (e.g., Heckman (2007)). A core focus of this literature has been to understand how susceptible children are to household-level shocks, variation in family resources, and changes in parental influences (e.g., Carneiro et al. (2022; 2021); Tungodden and Willen (2022); Willage and Willen (2022)). We contribute to this literature by providing the first quasi-experimental evidence on the intergenerational effect of sick leave - one of the largest social security programs in the OECD. This paper, therefore, helps advance our understanding of the interactions between existing social institutions and childhood development, and the importance of considering spillover effects across generations when designing basic social protection schemes.

Second, there is a small but rapidly expanding literature focusing on the intergenerational effects of specific welfare programs, such as disability insurance (Dahl and Gielen (2021); Dahl et al. (2014)) and U.S-specific anti-poverty programs AFDC, TANF, and EITC (Hartley et al. (2017)). We contribute to this literature by examining channels through which these effects may operate, exploring the impact of parental welfare usage on child educational outcomes, both on the intensive as well as the extensive margin.[4] In addition, the shocks we explore are less extreme, much more common, and intended as a temporary relief relative to programs such as disability insurance. Specifically, certified sick leave days make up the overwhelming majority of lost work days across the globe (Godoy and Dale-Olsen (2018)). Thus, sick leave is used by a significantly larger share of the population and is the first instance of

[4]In addition, our findings contribute to a long-standing debate on the intergenerational transmission of human capital and how to facilitate upward socioeconomic mobility (e.g. Black et al. (2005)).

employment protection against health challenges before workers have to resort to welfare programs of a more permanent nature.

Third, there is a small set of novel papers examining the impact of sick leave on individual workers (e.g., Fevang et al. (2014); Pichler and Ziebarth (2020); Markussen (2012); Godoy and Dale-Olsen (2018)). However, these studies do not have access to random variation in sick leave enrollment, and the examined outcomes have commonly been restricted to earnings and employment of the directly affected individual (i.e., the individual who is subject to the sick leave program).[5] We advance this literature by broadening the set of individual outcomes we explore and examining intergenerational spillovers. These results help us better understand the overall implication of sick leave on individuals and their children, and how that affects societal welfare.

More broadly, there is a wealth of observational studies on the effects of parental welfare utilization on children (see, for example, Black and Deveraux (2011), for an overview of these studies). However, many of these studies suffer from lack of exogenous variation in parental welfare usage, having to rely on fixed effects models with non-random variation in take-up. While a handful of studies have moved beyond the observational study design and exploited quasi-experimental variation, they have been forced to exploit variation across geography and time (e.g., Antel (2021); Levine and Zimmerman (1996)). Using conditional random assignment to GPs and exploiting a GP leniency design, we can overcome some of these challenges and provide carefully estimated and causally identified effects of parental welfare on children.

2 Background

In this section, we briefly provide an overview of the most relevant aspects of the Norwegian welfare state, the GP system, the laws governing patient reallocation across GPs, and the education system.

2.1 The Norwegian Welfare State and Paid Sick Leave

All permanent residents are automatically enrolled in the public social insurance system. This system is financed through a national insurance contribution imposed on both employers and employees. The system encompasses several welfare programs ranging from old age pension and health-related social insurance to transitional benefits for survivors and funeral grants. The three largest work-related social insurance programs are unemployment insurance (UI), sick leave, and disability insurance (DI). In terms of the degree of employment protection in Norway, the country is characterized as having a medium-to-high level

[5]The one exception is Godoy and Dale-Olsen (2018), which uses GP swaps in Norway to look at spillover effects of sick leave among colleagues at the workplace.

of protection relative to other OECD countries, on a level similar to Italy (Salvanes et al. (2023)).

The system for paid sick leave is designed to provide financial compensation for income loss caused by a temporary illness or injury. Sick leave benefits are paid by the employer for the first 16 days, and then by the government for a maximum of 52 weeks. The replacement rate is 100 percent from the first day of leave subject to a maximum amount ($62,000 per year in 2019). To qualify for sick leave benefits, an individual must have been employed for the past four weeks. Sick leave beyond three days requires a certificate from the worker's GP.[6]

After the sick leave period expires, individuals can apply for work assessment benefits, a time-limited extension to sick leave (but with benefits reduced from 100 to 66 percent) intended to provide support for rehabilitation and rest to facilitate reintegration into the labor market.[7] Should reintegration not succeed due to health-related challenges, the next step is often to apply for DI.

The two largest non-sick leave employment protection programs in Norway are UI and DI, and we explore spillover effects across these programs when studying the mechanisms behind our reduced-form effects. UI is available to all individuals who experience at least a 50 percent reduction in work hours and have a minimum income before becoming unemployed ($16,500 in 2019, Johnsen et al. (2022)). The replacement rate is approximately 62 percent, and the standard entitlement period is 104 weeks. Unemployment benefits are conditional on filing an employment form with the public employment office every 14 days, and on having a pre-dismissal income above a certain minimum threshold. The rules are more generous for older workers, and every worker over 60.5 is effectively entitled to UI until the mandatory retirement age of 67.

DI is provided to those who experience an injury or disability that causes a permanent reduction in earnings capacity. To receive DI benefits, a doctor appointed by the Labour and Welfare administration (NAV) must certify that the individual has attempted all appropriate treatments that could help improve their work ability. The DI replacement rate depends on an individual's pre-DI earnings. While the after-tax replacement rate can be above 100 percent for low-income groups, it decreases at higher incomes. The after-tax replacement rate for fully disabled, previously average earners, is around 65 percent.[8]

[6] In the public sector, workers can use eight days of sick leave before having to obtain a certificate from the GP. However, the effects are relatively similar across the public and the private sector (see Section 5 for results and discussion). If the injury is related to the musculoskeletal system, the individual can also obtain approval from a chiropractor or manual therapist. We abstract from this in the current analysis, something that may attenuate our results slightly.

[7] Before 2010, this was called rehabilitation benefits in the Norwegian system.

[8] More specifically, 66 percent of the three best years of the last five years leading up to program take-up.

2.2 The Norwegian GP System

The Norwegian health care system is a two-part system, with primary care provided by the local municipalities and specialist care provided by larger health regions. Access to specialist care and hospitals is only possible through referrals from GPs in the primary care sector (except in emergencies). The GP is, therefore, the first point of contact for non-emergency and preventive care, and is responsible for initial examination, diagnosis and treatment.[9]

In terms of the primary care system, every resident of Norway is assigned a general practitioner by the Norwegian Health Economics Association (part of the Norwegian Directorate of Health). In general, individuals must interact with their assigned GP every time they use the health care system. However, if the GP has already referred the patient to a specialist for a specific illness or problem, the patient may continue to use the specialist for that specific purpose without going through the GP. Individuals are allowed to change the GP they have been assigned twice a year conditional on availability (as described in Riise et al. (2022)).

GPs are traditionally self-employed, and municipalities contract with individual GPs to provide services to their local residents by assigning them a patient list.[10] GP earnings come primarily from fee-for-service from the health administration (around 70 percent), but also from capitation from the municipalities (approximately 30 percent) and out-of-pocket payments from patients (Ekspertutvalget (2023)).

To examine the intergenerational impact of sick leave through the use of a sick leave leniency measure, we require variation in GP assignment that is unrelated to patient characteristics. To this end, we exploit a novel feature of the law governing the Norwegian GP system (FOR-2012-08-29-842). When GPs retire, move, or for some other reason decide to terminate/reduce their current practice, the assignment of patients to new GPs is randomized by a computer system.[11] In the case of list reductions, neither the GP nor any other

See https://www.nav.no/uforetrygd/en

[9]Medical degrees are offered at seven universities across the country of Norway, all of which are public schools with identical curricula. Admission is based exclusively on GPA, and while it is slightly more difficult to get into the medical program at the University of Oslo, all seven schools have very similar GPA thresholds and are among the most difficult programs to get admitted to in the country. Thus, there is little variation in doctor training across the different institutions. See https://www.universitetsavisa.no/medisin-samordna-opptak-studentopptak/dette-studiet-har-na-702-i-poenggrense-ny-rekord/383387.

[10]In recent years, an increasing share of GPs have been hired directly by the municipality governments on a permanent contract. As of 2021, this share was 14 percent, see https://www.helsedirektoratet.no/rapporter/handlingsplan-for-allmennlegetjenesten-arsrapport -2022-inklusive-status-per-mai-2023/utvikling-fastlegeordningen

[11]Approximately 26 percent of all swaps in the Norwegian health care system are caused by doctors reducing or terminating their patient list, something we can observe directly in the data. The other two main sources of swaps are ordinary exchanges (65 percent) and automatic allocation (7 percent). Ordinary swaps refer to endogenous swaps initiated by patients (individuals are limited to perform two such swaps per calendar year), while automatic allocations refer to the assignment of patients to doctors the first time they enter the system (which would be at the time of birth for Norwegian-born individuals). We do not use

person has the right to decide who shall be transferred and who shall remain on the list (FOR-2012-08-29-842, Section 35). This also applies to the cases when a list is terminated but all patients are not transferred to the same new GP.[12]Furthermore, the computer system randomly assigns patients to new, available GPs, and this procedure is also the same for both terminations and reductions. The automated computer-based randomization procedure ensures exogenous variation in GP-patient match in our setting.[13]

When constructing our GP leniency measure, we do not use the initial assignments, nor any swaps initiated by the patients, due to endogeneity concerns. In addition, we always include pre-reassignment GP fixed effects, such that our leniency measures are identified by patients who initially had the same GP but then got exogenously allocated to new, and different, GPs. Thus, should a new GP take over the entire patient list of a retiring GP, those patients will not contribute to our identification. Importantly, each individual has the legal right to a primary GP at all times, such that there should be no point in time in which a patient lacks access to a GP.

The thought experiment underlying our empirical approach is to compare the children of parents who initially had the same GP (and thus lived in the same catchment area) and had the same baseline sick leave history, but who then were allocated to new, and different, GPs. The parents were allocated to these new and different GPs because the initial GP decided to reduce or terminate the patient list, and the Norwegian Health Economics Association randomly reallocated the patients across new available GPs. If all patients are sent to the same new GP, which could happen if a new GP takes over the entire list of an old GP, or if it is a rural area with only one GP, they do not contribute to identification (due to the inclusion of previous GP fixed effects).

2.3 The Norwegian Education System

The Norwegian education system encompasses ten years of mandatory education starting at age 6. The curriculum is set by the central government and the overwhelming majority of children attend public school (>95 percent).

Following the completion of compulsory education at grade 10, each student has the right to enroll in tuition-free high school (conditional on satisfactory graduation from compulsory school). The majority of Norwegian children pursue this option.

these swaps to identify sick leave leniency of GPs since they suffer from endogeneity issues.

[12]Information provided through private correspondence with the Norwegian Health Economics Association, which is governing this process.

[13]With respect to swaps generated by GP list reductions, these occur whenever a GP decides to reduce their patient list. While several different factors could trigger list reductions, a recent government-sponsored report alludes to increased workload, a change in the content and type of consultations that patients require, a change in work-life balance preferences of GPs, and a growing demand for availability and treatment intensity of patients, as being key drivers in inducing list reductions (see Ekspertutvalget (2023)).

High school in Norway (taken by students in grades 11 through 13 and correspond to students aged 16 through 18) consists of several program specializations within two types of tracks: an academic track and a vocational track. Students apply to high school through a centralized online system with the grades from their final year of compulsory education. The application consists of ranking three program specializations in the county of residence.[14] If the number of applications exceeds the number of available slots for a given program specialization, students will be assigned based on their grades in compulsory school.

High school education provides the student with university admission certification, vocational competence, or basic (craft) competence. University admission certification permits individuals to apply to, and enroll in, college. While university admission certification is awarded automatically to all students who successfully complete the academic high school track, individuals in the vocational track must take supplemental courses to attain this qualification.

Higher education is offered by a range of universities and colleges, the majority of which are tuition-free public institutions. Admission is coordinated through the Norwegian Universities and Colleges Admission Service. Students apply to specific programs at the different universities, and if the number of applications exceeds the number of available slots for a given program, students will be assigned exclusively based on their grades in high school.[15]

3 Data

The analysis performed in this paper requires linkages across several administrative data sets, and the data we use come from rich population-wide registers covering the universe of Norwegian residents and their health, education, and labor market histories. In terms of time period, we use exogenous swaps that occur between 2006 and 2018. Slight differences in the number of years for which we have access to the different outcome registers means that there will be small deviations in the number of observations across some of the analyses.

3.1 GP and Health Data

The Norwegian GP register provides information on the universe of all active GPs in the country for each year. Using unique GP identifiers, we combine this data with information from the *Control and Payment of Health Refunds Database (KUHR)*, which provides information on the number of times each patient has met the GP, the reason for the visit, the outcome of the visit and the total reimbursements for the GP. Importantly, these data also

[14] During our analysis period, Norway is divided into 19 administrative regions, called counties. The counties form the primary first-level subdivisions of Norway and are further divided into 431 municipalities. In 2020, the number of counties was reduced to 11, and the number of municipalities was reduced to 356. However, this does not coincide with our sample period.

[15] In addition, some programs impose specific course requirements such that only individuals who have taken certain high school courses are eligible for admission.

contain information on whether the GP has provided the patient with a sick leave certificate. The number of certified sick leave days actually taken by patients, and the diagnoses related to each specific sick leave spell, is merged in from *The Labour and Welfare Administration Database (NAV databasen).*[16]

Crucial to our analysis is the ability to identify the subset of GP-patient reallocations that are caused by existing GPs retiring, moving, reducing their practice, or exiting the market. In all cases where an entire list is not transferred to the same new GP, patients are randomized to new GPs through an automated computer system, as described above, thereby eliminating any risk of endogenous selection of particular patients to GPs. To identify these reallocations, we exploit the fact that the GP data also provide information on whether an individual changed GP during the year and the reason for that change. We focus on GP changes that are caused by the doctor deciding to terminate, or reduce, her patient list.

In Section 4.3, we provide evidence consistent with the notion that patient characteristics, including variables related to prior health care utilization, labor market outcomes, spousal characteristics, and baseline demographics, are uncorrelated with the sick leave leniency of the newly assigned GP. In Section 5.4, we also provide detailed evidence of GP leniency being unrelated to GP quality. Thus, the patient-doctor matches we exploit are plausibly exogenous, and the GP leniency measure we use is uncorrelated with other dimensions of GP care that could potentially have a confounding impact on the human capital development of the patients' children (through a health impact on the parents).

3.2 Child Education Data

Crucial to our analysis is the ability to link patients to their children, something we do through a unique family identifier. By following these children over time, from compulsory school into college, we can examine the impact of parental sick leave on children's short-and long-run education outcomes, both overall and as a function of the child's age at the time of parental GP change.

In terms of outcomes, we focus on a broad range of educational outcomes: GPA at the end of compulsory school (grade 10), high school GPA, the probability of pursuing an academic high school track, graduating from high school, and starting college. Summary statistics of these variables are provided in Panel A of Table 1. Taken together, these outcomes allow us to obtain a comprehensive understanding of the impact of sick leave leniency on children's educational outcomes in terms of performance, attainment, and behavior – both on the intensive as well as the extensive margin.

[16]Note that the number of sick leave days granted by the GP comes on top of the 3-8 sick days that can be taken without certification (3 if the individual works in the private sector and 8 if the individual works in the public sector).

3.3 Labor Market and Welfare Data

To better understand the mechanisms through which the child development effects operate, we follow parents across the administrative registers and collect key labor market and welfare information. These data are obtained from tax records governed by Statistics Norway and the social insurance database (FD-trygd), and provide detailed information not only on the employment status and wages of individuals, but also on all welfare programs they are enrolled in, how much benefits they receive from the various programs, and for what period. These data thus allow us to investigate if paid sick leave affects the employment, wage, and welfare dependence of the parent, and the extent to which these channels may help explain the child effects we observe.

In terms of outcomes, we begin by examining the effect of leniency on wages and employment. Our wage measure is based on pre-tax labor earnings (including income from self-employment) excluding government transfers. In addition to the employment and wage outcomes, we explore welfare dependence spillover effects to the main employment protection programs discussed above: DI and UI. In addition, we estimate the impact of leniency on the total amount of welfare transfers that the parent receives from the government. This provides us with a summary measure of the total impact of sick leave leniency on parental welfare usage. We estimate these effects not only in the short-run, but also five years after the swap took place. Summary statistics of these variables are provided in Panel B of Table 1.

In addition to the labor market and welfare participation mechanisms, we also examine the impact of sick leave leniency on the mental health of parents and their children. These data are taken from the GP and health data sources discussed above.

4 Method

4.1 Measure of sick leave leniency

Randomizing sick leave is not feasible: We cannot in practice force sick leave usage onto people. We can, however, think of a social experiment in which individuals are randomly assigned to doctors that differ in their willingness to provide sick leave certificates to their patients. This randomization would break the correlation between sick leave and unobserved determinants of individual labor market outcomes and intergenerational human capital developments. Comparing the effects of individuals who are randomly assigned to differentially "lenient" GPs would give a reduced form estimate of the effect of being exposed to an easier sick leave process.

The intention of our research design is to mimic this hypothetical experiment. Our source of exogenous variation comes from a novel feature of the Norwegian health care law in which patients are conditionally randomly allocated to new GPs in the event their current GP closes down or significantly reduces their practice. As GPs vary in their leniency of providing sick leave certificates to their patients, this randomization process allows us to construct a measure of GP sick leave leniency that is unrelated to any patient characteristics that may impact the child human capital development outcomes we explore.

To estimate our measure of sick leave leniency, we restrict our sample to patients who were employed at the time of the exogenous GP swap, as sick leave certification is conditional on having worked for the four weeks leading up to the sick leave request. We estimate the following equation:

$$h_{ijkt} = \mu_j + \pi_k + \boldsymbol{\theta_{it}} + \varepsilon_{ijkt}, \tag{1}$$

where h_{ijkt} represents the number of sick days of patient i in the year after exogenous assignment to GP j from GP k at time t. This measure includes zeros and is, therefore, capturing both the intensive as well as the extensive margin. The vector $\boldsymbol{\theta_{it}}$ includes controls for year-at-swap, age-at-swap, sex, and sick leave before swap; π_k are pre-swap GP fixed effects; and μ_j represent the exogenously-assigned new GP fixed effects. The new GP fixed effects μ_j form the basis of our leniency measure.[17]

By including pre-swap GP fixed effects and controlling for sick leave usage before the exogenous swap in Equation 1, the leniency measure is identified by a set of patients who had the same initial GP (and therefore lived in the same catchment area) and the same sick leave usage in the year before the swap, but then were allocated to new, and different, GPs due to factors unrelated to their health characteristics and the newly-assigned GP's quality.[18] Thus, our results are only identified off of individual patients who truly were exposed to exogenous shifts in GPs.

To use our leniency measure to identify the intergenerational effects of sick leave, we require physicians to be connected to each other through the patients they treat. As originally demonstrated by Abowd, Kramarz, and Margolis (1999) for the firm fixed effects framework, this is necessary in order to circumvent multicollinearity between subsets of patients and GPs. In other words, the pre-swap GP fixed effects and the exogenously-assigned GP fixed effects are only separately identified within connected sets of GPs. These GPs would be connected by patients from each pre-swap GP having different exogenously-assigned GPs (which is part of the research design) and by exogenously-assigned GPs who receive patients from multiple

[17] Location fixed effects (e.g., catchment area or municipality) are perfectly accounted for by the previous GP fixed effects (since the GP office location represents a finer level of geographic division).

[18] As discussed in Section 2, a new GP may take over the entire list of a retiring GP. This could be the case if there is only one GP present in a rural area or if a new GP takes over the entire list of a retiring GP. Those patients would not contribute to our identification due to the inclusion of previous GP fixed effects.

pre-swap GPs. To this end, we restrict our analyses to the largest connected group (this group includes 99 percent of all patients in our sample). To define our connected sets of GPs, we use all patients involved in exogenous GP re-assignments.

After calculating GP leniency, we construct a continuous standardized (mean 0, SD 1) measure of GP leniency $LeniencySD_j$ based on μ_j from above. In the robustness section, we also show results from a leave-one-out specification of μ_j, in which we exclude individual i from the leniency calculation when examining the impact of leniency on individual i's outcomes (e.g., Chetty et al. (2014b); Ginja et al. (2024); Jackson et al. (2020); Currie and Zhang (2023)).

4.2 Estimating impact on children and parents

We leverage the standardized measure of GP leniency described above to examine the effect of parental sick leave on child human capital development (main research questions) as well as the effect of parental sick leave on own labor market and welfare outcomes (mechanisms investigation). We estimate versions of the following equations:

$$w_{ijkt} = \beta LeniencySD_j + \pi_k + \boldsymbol{\theta_{it}} + \varepsilon_{ijkt}, \tag{2}$$

$$y_{cjkt} = \gamma LeniencySD_j + \rho_k + \boldsymbol{\gamma_{ct}} + \epsilon_{cjkt}, \tag{3}$$

where Equation 2 corresponds to our analysis on parents (denoted i) and Equation 3 corresponds to our analysis on children (denoted c). The coefficients β and γ represent the effects of a 1 SD increase in leniency on parents and children and corresponds to the parent receiving slightly more than two weeks of *additional* paid sick leave (corresponding to approximately 5 percent of a full work year). All other variables are defined as in Equation 1. In the event that an individual experienced several different GP changes during the analysis period, we focus on the leniency of the first exogenous GP swap.[19]

While we only observe the children's outcomes once (e.g., GPA and high school graduation), we observe many of the parents' outcomes on an annual basis (e.g., employment status and earnings). This allows us to provide complementary evidence using a difference-in-differences framework, leveraging within patient changes over time around the exogenous swaps. To this end, we estimate the following equation:

$$w_{ijkt} = \beta_1 Post_t + \beta_2 LeniencySD_j + \beta_3 (Post_t * LeniencySD_j) + \pi_k + \boldsymbol{\theta_{it}} + \varepsilon_{ijkt}, \tag{4}$$

[19] Restricting the sample based on multiple shock experiences of patients would introduce endogeneity into the estimation framework.

where β_3 represents the effect of receiving a more lenient GP on outcomes post swap, controlling for any systematic difference across patients assigned to differently-lenient GPs over time. Using this approach, we can also estimate pre-treatment trends in parental outcomes across differently lenient GPs to ensure that parents assigned to more or less lenient GPs were not on divergent trends prior to assignment. To this end, we restrict the sample period to the four years leading up to the exogenous swap and estimate the following equation:

$$w_{ijkt} = \beta_1 RT_t + \beta_2 LeniencySD_j + \beta_3(RT_t * LeniencySD_j) + \pi_k + \boldsymbol{\theta_{it}} + \varepsilon_{ijkt}, \tag{5}$$

where RT_t represents a continuous measure of time 1 to 4 years before the exogenous swap, and β_3 tests for the existence of pre-treatment relative trends between patients assigned to more or less lenient GPs prior to assignment. We obtain statistically non-significant and economically small point estimates for all our outcomes.

4.3 Identifying assumptions

The validity of our estimation framework hinges on the assumption that there is no systematic sorting of exogenously-assigned patients to GPs as a function of GP leniency. In other words, patients who experienced the same GP exit (either because of GP retirement, GP switching jobs, GP moving, or GP list reduction) cannot be systematically sorted across new GPs based on patient characteristics that also are correlated with their outcomes. In theory, the validity of this assumption follows directly from the fact that the Norwegian Health Economics Administration randomly reassigns patients to new local GPs conditional on municipality and availability in the event of GP list terminations or reductions. In practice, we can provide suggestive evidence on the validity of this assumption by showing that the GP leniency measure is unrelated to characteristics of patients that may also independently predict their outcomes.

To this end, we conduct a balance test in which we regress our estimated GP leniency measure on a large set of observable patient characteristics determined prior to the swap, including variables on health care utilization, labor market outcomes, baseline demographics, spousal characteristics, welfare usage, and family structure. Results from this exercise are provided in Table 2. All coefficients are economically small and none of the estimates are statistically significant at conventional levels. In addition, we conduct an aggregate balance test in which we use leniency as the left hand side variable and regresses it on all of the pre-determined characteristics simultaneously. These results are provided in Table 3 and demonstrate that the set of 19 pre-determined characteristics that we use are not jointly significant in predicting the leniency of the newly-assigned GP. The results from our balance tests provide support for the legally-mandated conditional random reallocation of patients to GPs in the event of GP exits and patient list reductions.[20]

Apart from the extensive balance checks, we use Equation 4 to provide complementary evidence using a difference-in-differences framework for the parental outcomes, leveraging within patient changes over time around the exogenous swaps. In addition, we use Equation 5 to estimate pre-treatment trends in parental outcomes across differently lenient GPs to ensure that parents assigned to more or less lenient GPs were not on divergent trends prior to assignment. Further, we conduct placebo tests in which we examine the effect of exogenous parental reassignment to a more or less lenient GP when children are ages 21 to 25 on outcomes when children are under the age of 20; we combine our baseline model with a propensity score matching approach; we estimate a version in which we drop children who are exposed to the same exogenous GP swap as their parent; and we employ a shrinkage approach to account for potential sampling error. We discuss all of the results from these exercises in Section 5.4.

In addition to the several robustness checks and sensitivity analyses discussed above, one remaining concern may be that more lenient GPs are of a different quality than less lenient GPs. If so, part of the effects we identify could operate through GP quality rather than GP leniency (GP quality affecting the health of the parent which could indirectly spill over to the child's human capital development). This would affect the interpretation of our findings. To address this concern, we analyse the relationship between GP leniency and (1) short-and long-term mortality at the patient level, (2) other GP practice characteristics at the doctor level, (3) GP value-added,[21], (4) GP reimbursements per visit as an indicator of treatment intensity and (5) inpatient visits, ER visits, and the likelihood that the GP conducts check-ups with the patient. Taken together, we find no relationship between GP sick leave leniency and this set of proxy variables for GP practice quality. This suggests that the GP behaviors underlying sick leave certification decisions are unrelated to their ability to improve patient health; a result consistent with prior research on the Norwegian sick leave and GP system (e.g., Markussen (2012); Markussen and Røed (2017). We discuss these results in detail in Section 5.4.

Finally, it should be noted that it is theoretically possible that individuals become aware of their doctor's impending retirement/move and switch to another doctor in anticipation.

[20] We also follow prior literature in examining the correlation of leniency across observable subgroups (e.g., Bhuller et al. (2018); Dobbie et al. (2018)). This approach provides a method for suggestively assessing the average monotonicity assumption required for causal inference in this literature. The results are shown in Appendix Figures A-1 and A-2, and the positive correlation across these subgroups provide strong suggestive evidence in favor of the required assumption.

[21] GP value-added is the 2-year post-assignment mortality of a GP's patients based on the conditional random assignment that we use for identifying leniency (see Ginja et al. (2024)).

Such behavior has no effect on the internal validity of our research design, because it does not have an impact on the randomization; patients who are exposed to an exogenous switch are still randomly reallocated to new GPs. However, it would have an impact on the external validity of our design, since it would influence which individuals are subject to randomization. It is, therefore, useful to understand whether this behavior exists in terms of describing who stays with a doctor until their retirement/move. To this end, we estimate the probability of endogenously switching GP as a function of the original GP retiring/moving/reducing the patient list in the next year. The results from this exercise are provided in Appendix Table A-1, and show a statistically significant but not economically meaningful effect: being exposed to an exogenous swap in the next year increases your likelihood of endogenously switching GP this year by 0.0005 (that we obtain a statistically significant estimate is perhaps unsurprising given the large sample size of more than 68,000,000 patient-year observations when performing this analysis).[22]

5 Results

5.1 Preliminary evidence on GP leniency

As noted above, our GP leniency measure is computed using exogenous reallocations of patients to GPs. In our data, there are 5,790 unique new GPs for whom we can calculate a GP leniency measure. Figure 1 shows where in the country these GPs are based. There is a higher density of new GPs in the populous cities of the country, and fewer new GPs in the northern part of the country where population density is low. On a per capita basis, the new GPs are uniformly distributed across the country.

Figure 2 shows the variation in sick leave duration (Panel A) and GP leniency (Panel B) in the year following the exogenous assignment of patients to new GPs. Panel A illustrates that approximately 15 percent of our sample experiences a paid sick leave spell during that first year, and that there is substantial variation in the duration of paid sick leave conditional on receiving sick leave.

The median sick leave spell, conditional on taking some sick leave, is around 90 days. Although shorter spells are more common, there is a substantial fraction of people who experience longer sick leave spells as well. For example, about 20 percent of individuals on sick leave experience between 90 and 180 days of paid sick leave, and 10 percent of individuals on sick leave experience between 180 and 270 days of paid sick leave. In addition, we see a non-trivial share of individuals bunching at the right-tail of the distribution (365 days); the maximum number of sick leave days an individual can receive in a given year year. These

[22] This analysis uses every patient in every year in the country of Norway, while our main analysis only uses patients with an exogenous swap. This analysis, therefore, has 68 million patient-year observations while our main sample has around 350,000 observations (depending on outcome).

results are presented in Panel A of Figure 2.

The pre-standardized GP leniency measure, obtained through the estimation of μ_j in Equation 1, is shown in Panel B of Figure 2. The leniency measure approximates a normal distribution relatively closely, with a mean of 0.5 and a standard deviation of 11. The figure demonstrates that being assigned to a GP who is located 1 SD above the mean leniency generates an *additional* 4.5 weeks (or 22 days) of paid sick leave relative to being assigned to a GP who is located 1 SD below the mean leniency on top of what is already taken. This is a substantial amount, equivalent to more than a month of full-time employment, or 10 percent of a normal work year in the country (or a 25 percent difference relative to the mean).

Panel C of Figure 2 formally shows the first stage of our instrument on parents own sick leave take-up using Equation 2, illustrating that the leniency measure we construct is highly predictive of parents' sick leave take-up (with an F-statistic of approximately 1500). In terms of magnitude, we find that a one SD change in leniency generates an increase in sick leave take-up of approximately 12 days, or 2.5 weeks. This effect is similar to the descriptive evidence based on the raw data in Figure 2, and represents an increase of 65 percent relative to the mean. Note that, for the core analysis of our project, we report reduced-form results based on Equation 3 rather than re-scaling those results through an IV approach. We do this because our findings based on Equation 2 show that part of the effects we identify operate through parents' own response to sick leave take-up. The interpretation of the reduced-form results therefore is clearer.

It is worth noting that the variation in sick leave leniency is much more pronounced for hard-to-verify musculoskeletal and psychological causes (Figure 3). These are diagnoses for which GPs have more individual freedom to choose both the length of sick leave and whether to grant sick leave or not. Sick leave leniency variation is much less pronounced for causes with little room for subjective interpretations, such as blood, eye, ear, and urology, related conditions.[23]

5.2 Effects on human capital development of children

Overall effects. Our core results regarding the effect of sick leave leniency on child human capital development, obtained from estimation of Equation 2, are displayed in Table 4. The outcomes we examine are GPA at the end of compulsory school (grade 10), high school GPA, the probability of pursuing an academic high school track, graduating from high school, and starting college.

The result in column 1 shows that children whose parents are exposed to a more lenient GP experience a reduction in education performance in lower secondary school. This effect is both economically meaningful and strongly statistically significant. In terms of magni-

[23] Appendix Table A-2 provides a full list of leniency standard deviations by ICPC code.

tude, the point estimate implies that children whose parents are exposed to a 1 SD more lenient GP experience a reduction in lower secondary GPA by 2.2 percent of a standard deviation. This performance effect is relatively sizable and is likely to have implications for the children's labor market outcomes; especially in light of recent evidence connecting small GPA changes to large differences in employer's hiring interest (Kessler et al. (2019)) and call-back rates (Quadlin (2018)). At the same time, these effects are smaller than the effects of targeted education inventions such as class size reductions or teacher quality improvements (e.g., Fredriksson et al. (2013); Chetty et al. (2014a); Krueger and Whitmore (2001)), and smaller than severe parental shocks on child human capital development such as job loss (e.g., Carneiro et al. (2022)).

In column 2, we examine the performance effect in upper secondary school. The point estimate in column 2 is very similar to that in column 1. The consistent performance effect across the different educational levels implies that the GPA effect identified in column 1 is not a short-term transitory effect, but likely a long-term permanent implication of GP sick leave leniency.

The results in columns 1 and 2 are important for disentangling the theoretical ambiguity surrounding the impact of employment protection on child human capital. As noted in Section 1, employment protection take-up could benefit child development by providing financial security, reducing stress, and improving the home environment; thereby increasing the probability of high-quality child-parent interactions. However, these programs may also hurt child development. This would be the case if program participation generates increased welfare dependence and worsened mental health, negatively impacts the parent's long-run labor market trajectory, and transmits negative role model/perception signals to the child. The results in columns 1 and 2 of Table 4 imply that the sum of the negative effects outweigh the positive, helping us to better understand the potential channels through which the intergenerational welfare dependence effects that have been identified in the existing literature operate (e.g., Dahl and Gielen (2021); Dahl et al. (2014); Hartley et al. (2017)).

In addition to affecting the intensive margin of educational performance, parental employment protection take-up could impact both the quantity as well as the quality of the human capital investments that the children undertake. To this end, we also explore the impact on the type and quantity of education in high school (columns 3 and 4 of Table 4) and college (columns 5 and 6 of Table 4).

In terms of high school effects, columns 3 and 4 demonstrate that GP leniency is associated with a decline in the likelihood of graduating, but not with a change in the probability to select into the academic versus vocational track. This suggest that parental welfare take-up has an overall impact on the amount of human capital investment that children make, but not on the type of investment that they make (conditional on making those investments).

In terms of college effects, columns 5 and 6 show that GP leniency negatively impacts children's probability to enroll in college, and generates an overall decline in the years of education that the children complete. For example, in response to the parent being assigned a GP who is 1 SD more lenient, we find a reduction in the probability of attending college with 1 percentage point. Thus, the impact of parental employment protection on child development does not only affect the quality of the human capital that children accumulate, but also the quantity. These results are in line with the negative impact on high school graduation in column 4, as well as with the negative performance effects identified in columns 1 and 2.

Reassuringly, Appendix Table A-3 shows that the effects on child human capital development are largest for leniency based on hard-to-verify causes with the greatest variance in the leniency measure (musculoskeletal) and smallest (not significant) for leniency based on easier-to-verify causes with the smallest variance in the leniency measure (respiratory).

Taken together, the results displayed in Table 4 highlight that the trade-off between social protection and work incentives extends beyond the individual worker, showcases the relationship between existing social institutions and child development, and demonstrates another dimension of the home environment through which children's human capital is shaped. Next, we explore whether the timing of enrollment matters and if the effects differ by sex, socioeconomic status, and parental gender. Finally, we examine the mechanisms through which these effects operate.

Timing effects. The life-cycle approach to skill formation suggests that children's development may not only depend on how much investment occurs during their childhood, but also on its timing (e.g., Heckman (2007)). For example, is well established that shocks and investments occurring in early childhood have long lasting consequences on children's development (e.g., Carneiro and Heckman (2003); Grantham-McGregor et al. (2007); Almond and Currie (2010)), and there has been a strong emphasize on the importance of early childhood interventions to combat poverty and exclusion. At the same time, recent observational studies (e.g., Carneiro et al. (2021); Eshaghnia et al. (2022); Eshaghnia and Heckman (2023), as well as contemporaneous causal analyses (e.g., Carneiro et al. (2022)), have suggested that similar shocks occurring at later stages of childhood may have similar, or even larger, long-term impacts. To examine this question in detail, Table 5 provides evidence on the effect of GP leniency on a selection of short-run (GPA) and long-run (college enrollment) child educational outcomes as a function of the age of the child at the time of the exogenous parental GP reassignment.

The results in Table 5 suggest that the timing of parental exposure to these programs appear to matter for how they impact child development. Specifically, exposure in early adolescence generally have larger impacts. However, even if this pattern in effect size is consistent with prior work on the timing of investment in children (e.g., Carneiro et al. (2021; 2022)), the effects are often not statistically significant across ages.

In terms of the current analysis, the timing results help us better understand the mechanisms underlying the child human capital effects. Specifically, these results imply that the effects on child development are unlikely to be driven exclusively by income effects of the parents (as this would imply larger effects for younger children who are exposed to these effects for a longer time), and at least partly driven by more short-run effects on components such as role model perceptions and stress.[24]

Heterogeneity effects. In light of recent literature documenting substantial effect heterogeneity in response to early childhood shocks across child sex, socioeconomic status, and parental sex, we perform a series of heterogeneity analyses to examine if certain children are more impacted by parental welfare take-up than others.

The results from this series of analyses are shown in Table 6 (parent sex), Table 7 (child sex), and Appendix Table A-4 (parental income). Overall, the heterogeneity analyses suggest that boys are more impacted by parental welfare take-up across a non-negligible share of the human capital measures we examine. We find no systematic differences across parent sex or the socioeconomic condition of the household.[25]

5.3 Effect on parental labor market, welfare, and mental health outcomes

After having examined the impact on child development, we proceed to investigate potential mechanisms that may explain these effects. We do this by linking the patient data to a rich set of labor market, welfare participation, and mental health outcomes. To avoid any mechanical relationship between the sick leave effect in the year following the swap and these parental outcomes, we focus on labor market and program substitution effects two years after the swap.

We begin by examining the impact on earnings and employment (Panel A of Table 8). The result in column 1 shows that sick leave certification has no effect on the employment prospects of the individual worker in the second post-swap year. In column 2, however, we

[24]Note that we have fewer observations for children who were of a very young age at the time of exposure (since we must wait at least until age 16 to collect outcome information on them). As such, it is problematic to split the sample into uniform age ranges (e.g., 3 year intervals). Instead, we have divided the sample into age groups such that the sample size is relatively stable across the groups while at the same time maintaining a meaningful age division. Because of this, the youngest age group encompasses many more ages, but still has a sample size that is slightly smaller.

[25]As noted in Section 2, in the public sector, the rules are slightly more generous then in the private sector. As such, we also conducted a stratified regression based on which sector the parent worked in. However, the effects are relatively similar across the public and the private sector (Appendix Table A-5). This suggests that the effects are not exclusively loading on one particular sector that has different types of rules.

show that the sick leave leniency generates a sizable drop in individual earnings. This cannot be a mechanical relationship caused by the government-mandated cap on sick leave benefits as the maximum spell length is one year (see discussion in Section 2), and must operate through a reduction in work quality, hours, or wage growth, as a direct implication of the leniency-induced sick leave spell. This result thus suggests a direct negative earnings effect of sick leave leniency, consistent with prior work on this topic (e.g., Markussen (2012)).

To examine if there are any program substitution effects into the two largest non-sick leave employment protection programs in Norway – UI or DI – we estimate the effect of GP leniency on take-up of UI and DI. To capture the extensive as well as intensive level effects, we explore both the probability of taking up these programs as well as the amount of benefits received through these programs. To ensure that we do not overlook any potential welfare dependence effect, we also examine the impact on any benefit take-up as well as the total benefit amount received from the national government.

Panel B of Table 8 shows an increased probability of receiving any benefits, an increased probability of receiving unemployment benefits (which must operate through a reduction on the intensive margin of employment given the lack of an extensive margin employment effect above), and no effect on disability usage. In terms of levels, we see a substantial increase in the total value of welfare benefits received and we do see suggestive evidence of positive increases on the intensive margin of both DI and UI (though the UI intensive margin effect is only marginally significant at the ten percent level, and the DI intensive margin effect falls just outside of the 10 percent significance level). The magnitude of the total welfare benefit effect is not sufficiently large to completely offset the income loss shown in column 2 of Panel A. Specifically, the result in column 5 of Panel B suggest that the increased welfare usage can mute approximately 45 percent of the income loss caused by GP leniency.

To understand the permanency of the effects in Table 8, we examine the impact on earnings and total welfare benefits five years after the exogenous swaps took place. These reduced-form impacts should be interpreted as the sum total long-run effect of GP sick leave leniency on earnings and welfare dependence. We believe that such aggregate five year effects are important for thinking about the parental welfare dependence and employment pathway as an explanation for the intergenerational sick leave effects we identify, especially since some of the child outcomes are measured several years after the shock took place.

Results are provided in Table 9 and show a clear long-run decline in earnings and a sizable increase in the use of the social insurance system. We hypothesize that the aggregation of these effects, in combination with any potential negative role model and perception effects, are driving the effect on the human capital development of children. As shown above, some of the human capital effects are larger if the parental GP reassignment takes place when the child is older rather than younger. Differences across child age are often not statistically significant, but the results suggest that the human capital effects are not exclusively driven by income effects (as this would imply larger effects for younger children who are exposed to these effects for a longer period of time). The effects must therefore, at least partly, also be driven by more short-run effects on components such as role model perceptions and stress.

While it is difficult to explore role model effects with our administrative data, we can examine stress and mental health effects through the patient registers described in Section 3. To this end, we provide results from estimating our main regression using psychological symptoms and diagnoses (column 1), anxiety and depression related symptoms and diagnoses (column 2), and substance abuse symptoms and diagnoses (column 3), in Table 10. Across all columns, we find economically meaningful and statistically significant effects: at the 5 percent level in column 1, at the 10 percent level in column 2, and at the 1 percent level in column 3. This set of results is consistent with negative mental health effects of welfare dependence and income loss that gradually develop over time and aligns with prior work on this topic (e.g., Kuhn et al. (2009); Carneiro et al. (2022)). These results are also consistent with recent work in child psychology and economics that documents a strong relationship between parental stress, parenting behaviors, and children's cognitive and socio-emotional skills (e.g., Yeung et al. (2002); Doepke and Zilibotti (2017); Carneiro et al. (2022)). While we are unable to isolate the extent to which these mental health findings drive the intergenerational impact of sick leave on child human capital development, it provides evidence on another suggestive channel through which our results may operate, and emphasizes the importance of additional research on this topic in future work. Note that we find no noticeable direct effects on the mental health of children (Appendix Table A-6). The human capital effects we find are therefore not operating through a direct mental health impact on the children, but plausibly – at least partly – through the impact that deteriorating mental health of the parent has on the interaction and relationship with the child.

5.4 Threats to identification

To ensure that our results are not driven by particular features of our research design, we conduct a series of robustness and sensitivity analyses on our main findings. These exercises can largely be divided into four groups: those that serve to examine the robustness and sensitivity of our results to the choice of model specification, those that are intended as placebo tests to ensure that we are not picking up spurious correlations, those that rule out alternative explanations, and those that provide alternative estimation approaches to provide complementary evidence on causal identification. We discuss each of these four groups of exercises below.

Complementary evidence. While we only observe child outcomes once, many of the parental outcomes we have information on are collected on an annual basis (e.g., employment and earnings). This enables us to provide complementary evidence using a difference-in-differences framework as specified in Equation 4, leveraging within patient changes over time around the exogenous swaps. Using this approach, we can also estimate pre-treatment trends in parental outcomes across differently lenient GPs to ensure that parents assigned to more or less lenient GPs were not on divergent trends prior to assignment; something that we do through the estimation of Equation 5.

The results from these exercises are provided in Tables 11 and 12. These results illustrate that we obtain quantitatively similar results on parental outcomes when using the alternative difference-in-difference strategy in which we only leverage within patient changes over time around the exogenous swaps (Equation 4). The results further illustrate that parents assigned to more or less lenient GPs are not trending differently in the years leading up to the swap; all pre-treatment trend differences are tightly estimated zeros. These results provide additional support for the exogeneity of the GP reassignments.

Placebo tests. Children who are older than the age at which the final outcome is measured should not be impacted by parental assignment to a more or less lenient GP. To this end, we also conduct placebo tests estimating the effect of exogenous parental reassignment to a more or less lenient GP when children are age 21 to 25 on outcomes that occur before age 20 (high school outcomes). The results from this exercise are provided in Table 13.

None of the results from this exercise are statistically significant even at the ten percent level, and all the coefficient for which we find significant effects in our main regressions (Table 4) are at least 60 percent smaller in these placebo regressions. This suggests that our main findings are unlikely to be driven by spurious correlations or endogeneity issues associated with unobserved selective sorting of parents to GPs with different sick leave leniency.

Alternative pathways. In addition to the placebo tests, supplemental analyses, and complementary evidence discussed above, one concern may be that more lenient GPs are of a different quality than less lenient GPs, have different practice styles, or cause their patients to use different amounts of healthcare. If so, some of the effects we identify could operate through aspects of the GP rather than GP sick leave leniency.

To examine this issue, we analyse the relationship between GP leniency and other potential pathways. First, we examine short-and long-term mortality at the patient level (Table 14). Second, we consider other GP practice characteristics at the doctor level (columns 2-6 in Table 15). Third, we consider GP reimbursements per visit as an indicator of treatment intensity (column 7 in Table 15). Fourth, we construct a GP value-added measure and relate this measure to GP sick leave leniency (column 1 in Table 15).[26] Finally, we study health

[26]GP value-added is the 2-year post-assignment mortality of a GP's patients based on the conditional

care utilization including inpatient visits, ER visits, and the likelihood that the GP conducts check-ups with the patient (Table 16). One of the characteristics we examine in Table 15 is list length. Examining list length is interesting because not only could it be an indication of doctor quality (as measured by the demand for the doctor's services), but it could also be the case that less busy doctors allow for fewer sick leave spells and more frequently ask patients to come back and extent sick leave, while very busy doctors may simply provide an extended period of sick leave from the start.

Across Tables 14 through 16, we find no relationship between GP sick leave leniency and the GP practice style, treatment intensity, or quality measures that we construct, nor any impact on patients' use of healthcare. This suggests that the GP behaviors underlying sick leave certification decisions are likely unrelated to their ability to improve patient health or other aspects of the doctor-patient dynamic. This is not too surprising given the subjectivity involved in the sick leave certification process – especially for the hard to diagnose conditions that drive much of our leniency variation. This subjectivity has also been identified in prior Norwegian studies where no relationship has been found between sick leave leniency and GP quality (e.g., Markussen (2012); Markussen and Røed (2017)). These results, in combination with the exogenous assignment mechanism governed by national law and the rich set of fixed effects in our main specification, make it unlikely that the effects we identify are confounded by a strong covariance between sick leave leniency and GP quality.

Robustness and sensitivity. In addition to the above exercises, we also adjust the sample restrictions and model specification to examine the sensitivity of our results to model alterations. The results from these exercises are provided in Table 17. To facilitate the interpretation of the results, Panel A contain our core results as a reference point.

First, one challenge with estimating μ_j is sampling error because each GP has a different number of patients for which we can calculate leniency. It is perhaps less of a concern in our setting given the number of patients per GP, but it may still generate non-negligible variation in the degree of certainty associated with leniency across GPs. To examine if this has an impact on our results, we follow Chetty et al. (2014b) and construct a Bayesian empirical estimator by adjusting the estimated leniency.[27] The results from this analysis are provided in Panel B of Table 17. The effects become slightly larger in magnitude after adjusting for potential sampling error, but provide strong support for our core results discussed above.

Second, a concern with using exogenous GP swaps is that children may swap GP at the

random assignment that we use for identifying leniency (see Ginja et al. (2024)).

[27]Specifically, we estimate $BE_j = \lambda_j Leniency_j$, where the shrinkage factor is $\lambda_j = \sigma_\mu^2/(\sigma_u^2 + \sigma_\epsilon^2/\eta_j)$. The term σ_u^2 represents the between-GP variation in the given outcome and σ_ϵ^2 is the within-GP variance in the given outcome. In other words, we take advantage of the fact that we observe the full load of patients for a GP in order to account for potential sampling error.

same time as the parent. In such cases, the effects we identify on child development could be driven by the direct impact of the GP leniency on the child, rather than through the effect of the GP leniency on the parent. To this end, we estimate Equation 2 using only children who do no experience the same exogenous swap as the affected parent. The results from this analysis are provided in Panel C of Table 17. Our results are unaffected by this adjustment.

Third, we note that parents who have not used sick leave earlier in their career are less likely to request sick leave certifications from new GPs than parents who have used sick leave earlier in their career. The presence of such never-takers may attenuate our results. To this end, we estimate Equation 2 using only parents who had been taking some type of sick leave in the year before the swap. This allows us to zoom in on the individuals that we believe are more likely affected by the leniency of the GPs that they are assigned. The results from this analysis are provided in Panel D of Table 17. Most of the point estimates become larger than our baseline results, but the main take-away from the analysis remain unaffected.

Fourth, we combine our baseline model with a propensity score matching approach. The rationale underlying this exercise is that we would like to obtain a treatment and control group that are as comparable as possible, to ensure a meaningful interpretation of the results. By combining our baseline model with a propensity score matching approach (in which we regress the probability of assignment to a GP with above median leniency as a function of baseline sick leave, year of swap, age at swap and gender), we avoid the risk of the estimates being driven by control and treatment units that are very different from one another and have little overlap in terms of background characteristics. Of course, the conditional random variation in GP assignment – regulated by Norwegian law – suggest that such differences should not exist. However, the propensity score matching approach provides us with an additional level through which we can examine this. The results from this exercise are provided in Panel E of Table 17, and demonstrate that our estimates are not statistically significantly different if we combine our baseline model with a propensity score matching approach, suggesting that the main effects are not identified off of control and treatment units that are very different from one another on observable dimensions.

Fifth, we estimate our results using a leave-one-out specification of μ_j, in which we exclude individual i from the leniency calculation when examining the impact of leniency on individual i's outcomes (e.g., Chetty et al. (2014b); Ginja et al. (2024); Jackson et al. (2020); Currie and Zhang (2023)). The results from this exercise are shown in Panel F and illustrate that the effects mechanically become a bit smaller when using this method, but that it has no impact on the pattern of results or the conclusions we draw from the analysis.

Sixth, we re-estimate our leniency measure only using exogenous swaps of non-parents and then use this measure to perform our main analysis. The idea behind this approach is akin to a split-sample approach in which we only use individuals that are not in our main analysis (non-parents) to calculate leniency and then use this leniency measure on those who are in the main analysis (parents). The results from this exercise are shown in Panel G and illustrate that the effects become a bit smaller when using this method, but that it has no impact on the pattern of results or the conclusions we draw from the analysis.

In addition to the results in Table 17, we also show in Appendix Table A-7 that our results are robust to removing baseline sick leave as a control (Panel B), and to removing baseline sick leave as well age at swap and parent gender fixed effects (Panel C); covariates included in the core regressions in order to improve precision and reduce noise. These results are encouraging, and support the randomization assumption on which the analysis rests.

Finally, we have estimated our main equation, sequentially eliminating specific counties and years from the analysis. The idea behind this exercise is to ensure that our results are not driven by a particular year or region of the country. The results from these analyses are provided in Appendix Figures A-3 and A-4. These figures suggest that the results are not driven by particular regions or years.

6 Conclusion

Children are highly susceptible to their home environments, and a rich literature has demonstrated how the structure and resources of families influence the human capital development of children. Employment protection and social security programs that shield children from abrupt changes to the home environment may therefore play an important role in their human capital advancement.

This paper uses legally-mandated conditional random assignment of patients to GPs to calculate a leniency measure of paid sick leave certification. We link these data to information on the human capital development of the patients' children. We find sizable negative effects of parental sick leave leniency exposure on the child's human capital development. In addition, we show that the timing of exposure may matter. In terms of mechanisms, we find that sick leave induces parents to earn lower wages, become more dependent on the social safety net, and experience negative mental health effects. We argue that part of the effects we identify on child human capital development likely operates through these channels.

The main contribution of this paper is to exploit exogenous variation in parental take-up of a key employment protection program that accounts for the overwhelming majority of lost work days across the globe and leverage rich register data to identify its effect on the human capital development of children across their childhood.

The results from this analysis have important policy implications. First, the results highlight that the trade-off between social protection and work incentives extends beyond the individual worker. Second, it showcases the relationship between existing social institutions and child development, and highlights another dimension of the home environment through which children's human capital is shaped. Third, it implies that the costs of these programs may be considerably larger than previously thought.

References

Almond, Douglas, and Janet Currie. 2010. "Human Capital Development before Age Five. NBER Working Paper No. 15827.." *National Bureau of Economic Research.*

Antel, J. 2021. "The intergenerational transfer of welfare dependency: Some statistical evidence." *Review of Economics and Statistics*, 74(3): 467–473.

Bhuller, Manudeep, Gordon B Dahl, Katrine V Løken, and Magne Mogstad. 2018. "Intergenerational effects of incarceration." In *AEA Papers and Proceedings.* 108 234–240, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Black, Sandra, and Paul Deveraux. 2011. "Recent Developments in Intergenerational Mobility." *Handbook of Labor Economics 4B.*

Black, Sandra, Paul Deveraux, and Kjell Salvanes. 2005. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *American Economic Review*, 95(1): 437–449.

Carneiro, Pedro, Italo Garcia, Kjell Salvanes, and Emma Tominey. 2021. "Intergenerational Mobility and the Timing of Parental Income." *Journal of Political Economy*, forthcoming.

Carneiro, Pedro Manuel, and James J Heckman. 2003. "Human capital policy."

Carneiro, Pedro, Kjell Salvanes, Barton Willage, and Alexander Willen. 2022. "The timing of parental job displacement, child development and family adjustment." *IZA Discussion paper*, 15630.

Chetty, Raj, John Friedman, and Jonah Rockoff. 2014a. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review*, 104 2593–2632.

Chetty, Raj, John Friedman, and Jonah Rockoff. 2014b. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economics Review*, 104(9): 2633–2679.

Cunha, Flavio, James Heckman, Lance Lochner, and Dimitriy Masterow. 2006. "Interpreting the evidence on life cycle skill formation." *Handbook of the Economics of Education (eds. Eric Hanushek and Finish Welch.*

Currie, Janet, and Jonathan Zhang. 2023. "Doing More with Less: Predicting Primary Care Provider Effectiveness." *Review of Economics and Statistics.*

Dahl, Gordon, and Anne Gielen. 2021. "Intergenerational Spillovers in Disability Insurance." *American Economic Journal: Applied Economics*, 13(2): 116–150.

Dahl, Gordon, Andreas Kostol, and Magne Mogstad. 2014. "Family welfare cultures." *Quarterly Journal of Economics*, 129 1711–1752.

Dahl, Gordon, and Lance Lochner. 2012. "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." *American Economic Review.*

Dobbie, Will, Jacob Goldin, and Crystal S Yang. 2018. "The effects of pre-trial detention on conviction, future crime, and employment: Evidence from randomly assigned judges." *American Economic Review*, 108(2): 201–240.

Doepke, Matthias, and Fabrizio Zilibotti. 2017. "Parenting with style: Altruism and paternalism in intergenerational preference transmission." *Econometrica*, 85(5): 1331–1371.

Ekspertutvalget. 2023. "Gjennomgang av allmennlegetjenesten. Ekspertutvalgets rapport (Expert Committee on the General Practitioner Service)."Technical report, Ministry of Health and Care Services.

Eshaghnia, Sadegh, and James J Heckman. 2023. "Intergenerational transmission of inequality: Maternal endowments, investments, and birth outcomes." *National Bureau of Economic Research.*

Eshaghnia, Sadegh, James J Heckman, Rasmus Landersø, and Rafeh Qureshi. 2022. "Intergenerational transmission of family influence." *National Bureau of Economic Research.*

Fevang, Elisabth, Simen Markussen, and Knut Roed. 2014. "The sick pay trap." *Journal of Labor Economics*, 32(2): 305–336.

Fredriksson, Peter, Bjorn Ockert, and Hessel Oosterbeek. 2013. "The long-term effect of class size." *The Quarterly Journal of Economics*, 128, p. 249–285.

Ginja, Rita, Julie Riise, Barton Willage, and Alexander Willen. 2024. "Does Your Doctor Matter? Doctor Quality and Patient Outcomes." *Journal of Political Economy Microeconomics.*

Godoy, Anna, and Harald Dale-Olsen. 2018. "Spillovers from gatekeeping - peer effects in absenteeism." *Journal of Public Economics*, 167 190–204.

Grantham-McGregor, Sally, Yin Bun Cheung, Santiago Cueto, Paul Glewwe, Linda Richter, and Barbara Strupp. 2007. "Developmental potential in the first 5 years for children in developing countries." *The lancet*, 369(9555): 60–70.

Hartley, R., C. Lamarch, and J. Ziliak. 2017. "Welfare reform and the intergenerational transmission of dependence." *IZA Discussion Paper No 10942.*

Heckman, James. 2007. "The economics. technology, and neuroscience of human capability formation." *PNAS*, 104(33): 13250–13255.

Heckman, James, and Stefano Mosso. 2014. "The economics of human development and social mobility." *Annual Review of Economics*, 6 689–733.

Jackson, Kirabo, Shanette Porter, John Easton, Alyssa Blanchard, and Sebastian Kiguel. 2020. "School Effects on Socioemotional Development, School-Based Arrests, and Education Attainment." *American Economic Review: Insights*, 2 491–508.

Johnsen, Julian, Kjell Vaage, and Alexander Willen. 2022. "Interactions in Public Policies: Spousal Responses and Program Spillovers of Welfare Reforms." *The Economic Journal*, 132(642): 834–864.

Kessler, Judd B, Corinne Low, and Colin D Sullivan. 2019. "Incentivized resume rating: Eliciting employer preferences without deception." *American Economic Review*, 109(11): 3713–44.

Knudsen, Eric, James Heckman, Judy Cameron, and Jack Shonkoff. 2006. "Economic, neurobiological, and behavioral perspectives on building America's future workforce." *PNAS*, 27(103): 10155–10162.

Krueger, Alan, and Diane Whitmore. 2001. "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR." *The Economic Journal*, 111 1–28.

Kuhn, Andreas, Rafael Lalive, and Josef Zweimüller. 2009. "The public health costs of job loss." *Journal of health economics*, 28(6): 1099–1115.

Levine, P., and D. Zimmerman. 1996. "The intergenerational correlation in AFDC participation: Welfare trap or poverty trap?" *Institute for Research on Poverty Discussion Paper 1100-96.*

Markussen, Simen. 2012. "The individual cost of sick leave." *Journal of Population Economics*, 25(4): 1287–1306.

Markussen, Simen, and Knut Røed. 2017. "The market for paid sick leave." *Journal of health economics*, 55 244–261.

Pichler, Stefan, and Nicolas Ziebarth. 2020. "Labor Market Effects of U.S. Sick Pay Mandates." *Journal of Human Resources*, 55(2): 611–659.

Quadlin, Natasha. 2018. "The mark of a woman's record: Gender and academic performance in hiring." *American Sociological Review*, 83(2): 331–360.

Riise, Julie, Barton Willage, and Alexander Willen. 2022. "Can Female Doctors Cure the Gender STEMM Gap? Evidence from Exogenously Assigned General Practitioners." *Review of Economics and Statistics*, 104(4): 621–635.

Salvanes, Kjell, Barton Willage, and Alexander Willen. 2023. "The Effect of Labor Market Shocks Across the Life Cycle." *Journal of Labor Economics.*

Tungodden, Jonas, and Alexander Willen. 2022. "When Parents Decide: Gender Differences in Competitiveness." *Journal of Political Economy*, forthcoming.

Willage, Barton, and Alexander Willen. 2022. "Postpartum Job Loss: Transitory Effect on Mothers, Long-run Damage to Children." *European Economic Review.*

Yeung, W Jean, Miriam R Linver, and Jeanne Brooks-Gunn. 2002. "How money matters for young children's development: Parental investment and family processes." *Child development*, 73(6): 1861–1879.