Universidad del Rosario | Facultad de Economía

SERIE DOCUMENTOS DE TRABAJO

No. 263

Octubre de 2020

Does Higher Education Reduce Mortality? Evidence from a Natural Experiment in Chile

María Angélica Bautista

Felipe González

Luis R. Martínez

Pablo Munoz

Mounu Prem

Universidad del Rosario, Facultad de Economía
Cl. 12c #4-59. Bogotá, D. C., Colombia
PBX: (031) 2970200 ext. 4163

facultadeconomia@urosario.edu.co
https://www.urosario.edu.co/Facultad-de-Economia/Inicio/
https://repository.urosario.edu.co/handle/10336/20291

Does Higher Education Reduce Mortality? Evidence from a Natural Experiment in Chile*

María Angélica Bautista Felipe González Luis R. Martínez
Pablo Muñoz Mounu Prem

October 2020

Abstract

We exploit the sharp downward kink in college enrollment experienced by cohorts reaching college age after the 1973 military coup in Chile to study the causal effect of higher education on mortality. Using micro-data from the vital statistics for 1994-2017, we document an upward kink in the age-adjusted yearly mortality rate among the affected cohorts. Leveraging the kink in college enrollment, we estimate a negative effect of college on mortality, which is larger for men, but also sizable for women. Intermediate labor market outcomes (e.g., labor force participation) explain 30% of the reduction in mortality. A similar upward kink in mortality over multiple time horizons is also present among hospitalized patients in the affected cohorts, with observable characteristics (i.e. diagnostic, hospital, insurance) explaining over 40%. Survey responses reveal that college substantially improves access to private health care, but has mixed effects on health behaviors.

*We thank Pablo Celhay, José Ignacio Cuesta, Felipe Menares, and Maria Perdomo-Meza for comments and suggestions. We also thank Fondecyt (project 11170258) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 for financial support. González: Pontificia Universidad Católica de Chile, Instituto de Economía. Muñoz: FGV EPGE. Maria Ballesteros, Leonor Castro, and Anita Zaldivar provided outstanding research assistance.

1 Introduction

The relationship between education and health ranks among the most widely studied in economics. Observational evidence suggests that more educated people live longer lives, but existing research on the causal effect of education on mortality has provided largely null findings (Galama et al., 2018). However, previous work has almost exclusively focused on secondary education, exploiting changes in compulsory schooling laws. It seems likely that the effect on mortality varies at different levels of education, insofar as they differentially affect socioeconomic outcomes and health behaviors (Cutler and Lleras-Muney, 2008; Montez et al., 2012). Little is known about the impact of higher education on mortality, even though this relationship is of particular interest amid ever-rising prices and ongoing debates about the sign of the pecuniary return to college and the optimal level of government funding.

In this paper, we provide new evidence from Chile on the causal effect of college education on health. For this purpose, we exploit plausibly exogenous variation in university enrollment among cohorts that reached college age in a narrow window around the 1973 military coup that brought General Augusto Pinochet to power. As documented by Bautista et al. (2020b), the Pinochet regime quickly assumed control of all universities in the country and steadily reduced public subsidies to higher education over the following years. This was part of the regime's technocratic reforms, aimed at reducing inefficient spending and fostering universities' financial independence. The result, however, was a continual reduction in the number of openings offered by universities for incoming college students and a sharp reduction in the college enrollment rate for men and women reaching college age during this period.

Our empirical strategy exploits the downward kink in college enrollment for the affected cohorts to study the link between higher education and mortality. Not only is death the ultimate health cost, but it has the additional feature of being an objective outcome that we can study through administrative sources. We uncover an opposite upward kink in the age-adjusted mortality rate among the affected cohorts and estimate a negative causal effect of college on mortality. This effect is larger for men, but also sizable for women. Labor market outcomes affected by college enrollment, such as labor force participation, explain a sizable share of the reduction in mortality. An upward kink in mortality over multiple time horizons is also present among hospitalized patients in the affected cohorts, with observable characteristics such as diagnostic and type of insurance explaining (in a statistical sense) almost one half. Survey responses reveal that college increases enrollment in private health insurance, but has mixed effects on health behaviors.

In the first part of the paper, we combine information from individual records in the 1992 population census and the vital statistics to calculate yearly mortality rates at the cohort-region-gender level for the period 1994-2017. In doing so, we exploit information on educational attainment available in both sources and restrict the sample to individuals with complete secondary education, to ensure a relevant counterfactual for college enrollment. We also restrict the sample to individuals reaching age 21 between 1964 and 1981, leaving us with an 18-cohort window centered around 1973. We then provide reduced-form estimates of changes in the mortality trend for the cohorts reaching college age after the military coup. Our preferred specification exploits the 24-year panel and includes year by region of residence and age fixed effects, thereby allowing mortality to vary flexibly at each point in the life cycle. Using the post-coup kink as an excluded instrument, we also provide instrumental variables (IV) estimates of the effect of college enrollment on mortality, in the spirit of a Regression Kink Design (Card et al., 2015). The large decline in enrollment experienced by the affected cohorts (42% decrease between 1972 and 1981) makes for a very strong first stage. Moreover, Bautista et al. (2020b) show that this decline had large negative socioeconomic consequences, increasing the plausibility of health effects.

The exclusion restriction for our IV estimates requires that any change in the mortality of the affected cohorts is purely driven by the reduction in college enrollment. Underlying this assumption is the idea that any other changes brought about by the military regime should have affected contiguous cohorts of young adults in a roughly similar fashion. In this regard, we show that our results are unaffected if we use tighter bandwidths (as little as four cohorts on each side). To the best of our knowledge, there were no other events that could have affected the health of the study cohorts differentially. Furthermore, secular improvements in health conditions that disproportionately benefit younger cohorts (e.g. increased awareness of the health risks of smoking) will arguably play against us, given that we study a reduction in educational attainment among younger cohorts. Also relevant is the finding by Bautista et al. (2020b), which we reproduce below, that the kink in college enrollment was entirely supply-driven, as the affected cohorts display no change in secondary completion and their college applications always exceeded the number of openings.[1] Further tests based on Conley et al. (2012) indicate that violations of the exclusion restriction would have to be quite large (70% or more of the reduced-form effects) to make our findings insignificant.

Our reduced-form estimates provide evidence of upward kinks (i.e. increases) in the mortality rate for men and women that reached college age in the years after the military coup. The corresponding IV estimates from our preferred specification with age fixed effects indicate that college enrollment decreases the yearly probability of death by 0.26 percentage points (pp) for women and by 0.92 pp for men. These are large and precise effects, equivalent to 7% and 13% of the

[1] Bautista et al. (2020b) also show that the socioeconomic composition of the student body and the distribution of students across fields of study were not affected.

respective sample means, and are significantly different from each other at the 0.1% level. The average person in our sample is 52 years old, so our results should be interpreted as capturing the effect of education on middle-age mortality. Reductions in deaths from cancer or from diseases of the circulatory or digestive systems represent 85% of the effect of college on female mortality and 74% of the effect on male mortality. These results are robust to different bandwidths in terms of (i) cohorts, (ii) years, or (iii) ages.

Our IV estimates correspond to a Local Average Treatment Effect (LATE) for the population of compliers whose college enrollment was negatively affected by the military coup. These are the missing college students that the technocrats guiding economic policy for the military regime argued were undeserving of educational subsidies. Importantly, the IV coefficients are only slightly larger and not statistically different from the partial correlations estimated through OLS. This suggests that either positive selection into college is limited or that the reduction in mortality caused by college enrollment was particularly high among our compliers (Card, 1999). It is also possible that our group-based (i.e. cohort-based) instrument is capturing spillover effects that the OLS estimate fails to incorporate (Grossman, 2006).

In the second part of the paper, we conduct three different sets of exercises using additional data from different sources to shed light on the underlying mechanisms. First, we exploit harmonized information on labor market outcomes (labor force participation, occupation, type of employment) available in both the 1992 census and the mortality files to study mortality within more-tightly defined categories based on these outcomes. We find that conditioning on broad labor market categories reduces the magnitude of the effect of college on mortality by around 30% for both men and women. This indicates that market mechanisms (i.e. income and occupation) play an important mediating role.

Secondly, we exploit the universe of hospital discharge summaries between 2002 and 2018 in two different ways. We first study the hospitalization rate as an intermediate health outcome that is affected both by underlying health conditions and by access to medical care. We find no evidence of a kink in total hospitalizations for the affected cohorts, which is consistent with worse health and reduced access to care largely offsetting each other. In line with this interpretation, once we disaggregate by type of insurance we find a positive kink in admissions using the public insurance (FONASA) and a negative kink in admissions using private ones (ISAPRE). We also combine the discharge summaries with the mortality files at the individual level to study the mortality of hospitalized patients over different time horizons. *Ex-ante*, it is unclear whether education affects mortality once someone is hospitalized. We find evidence of upward kinks in mortality for patients in the affected cohorts in both the short and the long term (i.e. as much as five years). Observable characteristics, including diagnostic, type of insurance, hospital, whether surgery was performed and whether the patient had been previously admitted, statistically explain over 40% of this kink.

Third, we use data from 13 waves of a large household survey (CASEN) between 1990 and 2017 to further study access to medical care and health behaviors. We find that college enrollment has a large negative effect on enrollment in the public health insurance. This suggests that reduced reliance on the more congested public health system contributes to the reduction in mortality caused by college. Consistently with this interpretation, our IV estimates show that college enrollment increases the probability of having recently seen a general practitioner or a specialist. Regarding health behaviors, we find that college enrollment has a large positive effect on the probability of smoking, which we interpret as universities being particularly liberal environments that tolerated or even fostered risky behaviors by young adults in our setting. At the same time, we show that college increases the probability that women have had a Pap smear in the past three years, which helps explain the observed reduction in mortality from cervical cancer.

Our paper complements a large literature studying the non-pecuniary effects of education (Grossman, 2006; Oreopoulos and Salvanes, 2011), particularly the strand focused on the causal effect of education on health (Galama et al., 2018). An extensive body of work dating back to Kitagawa and Hauser (1968) has documented an educational gradient in health status, but causal evidence remains limited, as plausibly exogenous sources of variation in educational attainment are hard to come by. Most of the existing research on mortality has exploited changes to compulsory schooling laws at lower levels of education and has found largely null results.[2] However, the findings from these studies are highly localized and have limited external validity to other points in the distribution of education, higher education in particular, which may have different effects on health behaviors and mediating socioeconomic outcomes (Cutler and Lleras-Muney, 2008, 2010). The existing literature has also largely focused on European countries with relatively equitable access to high-quality medical care.

To the best of our knowledge, only Buckles et al. (2016) have studied the causal effect of college on mortality, exploiting the avoidance of the military draft by young males in the US during the Vietnam war (Card and Lemieux, 2001).[3] They document a negative effect (i.e. reduced mortality), but face the complication that men in cohorts exposed to the draft were also more likely to participate in the war. Women and non-white men are excluded from their sample. We add to the literature by introducing a novel empirical strategy to estimate the causal effect of college enrollment on female and male mortality, both in the population at large and among hospitalized patients. Our setting further provides a unique opportunity to combine data from multiple sources to gain insight on the mediating role of labor market outcomes, consumption of health services and

[2] See Lleras-Muney (2005); Mazumder (2008); Albouy and Lequien (2009); Van Kippersluis et al. (2011); Clark and Royer (2013); Black et al. (2015); Meghir et al. (2018). Papers studying other health outcomes are similarly inconclusive (Arendt, 2005; Oreopoulos, 2006; Braakmann, 2011; Kemptner et al., 2011).

[3] Grimard and Parent (2007) and De Walque (2007) exploit the same source of variation to study the link between college education and smoking behavior. Contrary to us, they find that college reduces the probability of smoking.

other observable characteristics in the causal link between education and health.

2 Institutional Background

In this section, we provide an overview of higher education in Chile and the changes it experienced following the 1973 military coup. We then provide a brief socioeconomic characterization of the country and basic information on the Chilean health system.

2.1 Regime Change and Higher Education

President Salvador Allende was overthrown by a military coup on September 11, 1973. At the time, there were eight universities in existence in Chile, with campuses spread throughout the country. Only two universities were public, but the entire system depended almost exclusively on government funding (77% of total revenue in 1972). Differentiated tuition based on family income was charged, but fees were generally low. Starting in 1967, admissions were determined by a matching algorithm based on applicants' preferences, their score in a centralized admissions exam and the yearly number of openings that universities made available for each of their programs.

A military junta presided by General Augusto Pinochet assumed control of the government after the coup and would remain in power until 1990. The junta quickly appointed members of the armed forces as rectors to all universities and endowed them with full discretion over university administration. However, policy quickly begun to be dictated by a group of technocratic economists leading the regime's efforts at economic stabilization and modernization. These economists would come to be known as the "Chicago Boys", as most had studied at the University of Chicago and were strong supporters of the free-market views associated with the Chicago school (Valdés, 1995). The Chicago boys advocated for a reduction in public subsidies to universities, arguing that these were inefficient and failed to promote effort and thrift. Following their advice, the military government drastically cut back on its contributions to universities over the following years. Panel (a) in Figure 1 shows that the share of the government's education budget allocated to universities fell from a high of 47% in 1974 to 29% by 1980 (38% decrease).

The reduction in government transfers to universities was not matched by increases in tuition (which faced strong internal resistance) or in other sources of self-generated revenue. Universities were thus forced to scale down and continually reduce the number of openings for incoming students, as panel (b) in Figure 1 shows. This led to a dramatic reduction in the college enrollment rate, as a growing number of graduates from secondary education competed over a decreasing number of spots. Panel (a) shows that college enrollment had risen steadily in the decade before the coup, reaching a gross enrollment rate of 9% (77,000 students) by the end of the Christian-Democrat government of Eduardo Frei in 1970. The enrollment rate grew even more during the

Allende government and peaked at 17% (146,000 students) in 1973. It steadily declined after the coup and was back down to almost 10% in 1980.[4] Panel (b) also shows that the number of college applicants was higher than the number of openings in all years, which implies that the supply of openings was always the binding constraint determining the number of incoming college students.

Naturally, the initial period following the military take-over brought about other changes to universities. In particular, the first months after the coup were characterized by highly-targeted repression against supporters of the deposed Allende and other political activists, though "most previously enrolled students remained enrolled" (Levy, 1986, p. 101).[5] Around two dozen research centers and academic units deemed politically undesirable were shut down and extracurricular activities were banned or tightly controlled (Brunner, 1984). However, the vast majority of academic units continued functioning and the distribution of students across fields of study hardly changed (Bautista et al., 2020b). Moreover, several important features of the system were also left unchanged. For instance, the regime preserved the centralized admissions process, which remains largely the same until the present day. Hence, even though the initial wave of expulsions was highly targeted, the ensuing reduction in openings (which provides the bulk of the variation in college enrollment that we use in the empirical analysis below), was not.[6] Students with lower test scores (i.e. closer to the counterfactual admissions cut-off) were the ones that failed to gain admission. Bautista et al. (2020b) show that applicants from less affluent socio-economic backgrounds were disproportionately affected, but they represented a small share of enrollment and the socioeconomic composition of the student body was left largely unaffected. The number of universities and campuses would also remain unchanged until a large reform in 1981.

2.2 *Socioeconomic Characterization*

Chile has experienced rapid economic growth since the mid-1980s, with GDP per capita (in constant 2010 USD) rising from $4,700 in 1985 to $14,700 in 2015. These numbers respectively correspond to 0.7 and 1.5 times the average income per person in Latin America, giving rise to the notion of the Chilean 'economic miracle'. Chile has been a member of the OECD since 2010. However, the benefits of economic growth have not accrued to everyone equally and Chile's Gini index of 0.46 in 2014 was the highest among OECD countries, which averaged 0.32.

Chile's sustained economic growth over the past decades has been reflected in improved health outcomes. Life expectancy in 1970 was 62.3 years, only slightly higher than the Latin American

[4] The unexpected nature of the drop in enrollment is supported by the fact that UNESCO projected 200,000 university students for 1975, highly overestimating the real figure of 150,000 (Levy, 1986).

[5] At its most extreme, state repression during the Pinochet dictatorship led to the death or forced disappearance of about 3,200 people (Bautista et al., 2020a). Using detailed records on victims from Comisión Rettig (1996), Bautista et al. (2020b) estimate death rates of 0.2% for both university students and faculty.

[6] Bautista et al. (2020b) show that the negative kink in college enrollment after the military coup is visible among siblings (i.e. family fixed effects) and within quintiles of housing wealth in 1992.

average of 60.4, but rapidly increased over the following decades, catching up with the OECD average of 75 years by the early 1990s. The infant mortality rate plummeted over the same period, from 67 deaths per 1,000 live births in 1970 to 16 in 1990, slightly lower than the OECD average of 17.5. Chile is also near the OECD average in terms of avoidable mortality and self-rated health (OECD, 2019). However, the country has higher-than-average rates of smoking (25% of adults), obesity (74% of adults) and chronic disease morbidity (9% of adults with diabetes).

2.3 Health system

Health insurance in Chile operates under a dual system that includes several private providers and a public alternative via the National Health Fund (*Fondo Nacional de Salud - FONASA*). This system dates back to the period 1979-1981, and was created as part of the Chicago Boys' technocratic reforms.[7] FONASA is a pay-as-you-go system financed with government funds and a 7% payroll tax. It has no exclusions and offers three levels of copay (0, 10 or 20%) based on income and number of dependents. The private providers, known as ISAPREs (*Instituciones de Salud Previsional*), are insurance companies that compete by offering contracts at different prices in a regulated market.[8] ISAPREs receive the payroll tax contributions made by their members and usually require additional payment. In 2011, payments to ISAPREs averaged 10.3% of wages, while the average copay was close to 33% (Galetovic and Sanhueza, 2013).

The share of the population covered by FONASA has risen over time, from around 66% in the 1990s to almost 80% in more recent years, while the share affiliated to an ISAPRE is now close to 14%, falling from a maximum of almost 25% in the mid-1990s (MDS, 2018).[9] These changes are partly explained by the rising cost of private insurance, which increased in real terms by a factor of 2.2 between 1991 and 2011, while real wages only increased by a factor of 1.8 over the same time period (Galetovic and Sanhueza, 2013). Cream skimming in this market is well documented: FONASA serves lower-income and riskier people, while ISAPREs serve a richer, healthier, and younger segment of the population (Pardo and Schott, 2012). Switching between ISAPREs and FONASA is relatively uncommon, though FONASA often acts as a safety net and absorbs people that lose their job (Duarte, 2011).

Health service provision also involves both private and public providers: laboratories, clinics, hospitals. In 2016, 24% of the 348 hospitals in the country were private (Clínicas de Chile, 2016). ISAPREs tend to offer full flexibility over providers and reduced copay for in-network or preferred

[7] Before then, public health insurance was comprised of separate white-collar and blue-collar health funds (SERMENA and SNS), which were created in 1942 and 1952, respectively.

[8] These firms can implement risk pricing or risk selection based on gender and age. In 2012, ISAPREs made available more than 52,000 different health insurance plans (Galetovic and Sanhueza, 2013).

[9] The share without insurance has steadily fallen is now at around 3%. The small remaining share (< 4%) either gets separate insurance through the Armed Forces or is affiliated to some other private insurer.

providers. In 2012, 97% of payments made by ISAPREs went to private providers (Galetovic and Sanhueza, 2013). FONASA, on the other hand, mostly covers services by public providers and additional payment is required to access private providers. Public providers tend to be more crowded and have longer wait times, as more than half of the country's physicians (55%) work in the private sector (Clínicas de Chile, 2016). However, a thorough comparison of the quality of services provided by public and private providers is beyond the scope of this paper and is made difficult by underlying differences in the populations served.

Health spending as a percentage of GDP has risen over time, from around 5% in the 1990s to 9% in 2017, the OECD average. Still, Chile ranks highly in terms of the efficiency of health expenditure. For example, the Bloomberg Health-Care Efficiency Index ranks the country 8th among 55 considered (Clínicas de Chile, 2016). Private spending represented 56% of the total in 2015, 85% of which corresponded to out-of-pocket expenses.

3 Data

We rely on four main data sources for the analysis. First, individual death records from the vital statistics for the period 1994-2017. Secondly, individual records from the 1992 and 2002 population censuses. Third, the universe of hospital discharge summaries between 2002 and 2018. Lastly, individual responses from the CASEN household survey between 1990 and 2017. In the rest of this section, we introduce our criteria for inclusion in the sample, provide an overview of each of these sources and describe the construction of our main variables of interest.

Following Bautista et al. (2020b), we restrict the sample to individuals born between 1943 and 1960. These individuals reached college age (i.e. age 21) between 1964 and 1981, creating an 18-cohort window centered around 1973, the year of the military coup.[10] We end the sample with the 1981 cohort to mitigate the confounding effect of the large reform of the Chilean university system that was implemented by the military regime after that year. Starting with the 1964 cohort creates a balanced sample centered around 1973. The discrete nature of the running variable prevents us from applying a non-parametric approach to select an optimal bandwidth, but we verify that our results are robust to other bandwidths. To ensure a relevant counterfactual for college enrollment, we further restrict the sample to individuals reporting four or more years of secondary education (*educación media*).[11] Our results are also robust to removing this constraint, though their interpretation naturally changes, as we discuss below.

The Department of Health Statistics and Information (*Departamento de Estadísticas e Información de Salud - DEIS*) provides rich individual-level data from the death certificates. Basic

[10] Bautista et al. (2020b) show that the average age of first-year college students in 1970 was 20.5 years.

[11] Information on secondary completion is unavailable in these sources.

information on each deceased individual includes year of birth, gender, educational attainment, county of residence, and cause of death. The death records also include additional information on labor market outcomes (labor force status, occupation and type of employment) that we use in our analysis of mechanisms.[12] In order to calculate mortality rates, we use information from the 1992 census provided by the National Institute of Statistics (*Instituto Nacional de Estadística - INE*). Besides basic demographic characteristics, the census provides the same information on educational attainment and labor market outcomes as the death certificates.

To construct our main outcome of interest, we proceed as follows. For each year between 1994 and 2017, we calculate the risk-adjusted yearly mortality rate at the cohort-gender-region level.[13] In doing so, we follow Clark and Royer (2013) and iteratively adjust the initial population count in the 1992 census for the number of deaths per cell in the previous year. We initially observe 997,484 individuals that meet our sample criteria in the census, with ages between 32 and 49. The average individual in the panel is 52 years old. We observe 124,729 deaths among these individuals during the sample period, yielding an aggregate mortality rate of 12.5%. Male mortality is almost twice as high as female mortality (16% vs 9%), similarly to other settings (Beltrán-Sánchez et al., 2015).

Using the information on the cause of death reported in the mortality files, we replicate the previous procedure to obtain cause-specific mortality rates.[14] We also calculate the share of people in each cohort-gender-region cell that report any college education in the census and iteratively adjust this share based on the educational attainment of the people that passed away from that cell in the previous year. When we first observe the study cohorts in 1992, the youngest one has age 32, so we can confidently assume that people in the sample have completed their education. Our results are robust to using the unadjusted mortality rates and/or college shares.

The third piece of data comes from the universe of hospital discharge summaries between 2002 and 2018. This dataset has almost five million observations and reports basic demographic information of the patient as well as the hospital of admission, diagnostic and type of insurance, among other characteristics. Unfortunately, it does not include information on educational attainment, so when using this data we cannot restrict the sample to individuals with full secondary and can only provide reduced-form results. We combine this data with the 2002 census to construct yearly hospitalization rates at the cohort-gender-region level for the period 2002-2018. The discharge summaries have a unique individual identifier that allows us to distinguish between patients that are hospitalized for the first time and those that are readmitted. Approximately 37% of records

[12]Though available, the mortality files for the period 1990-1993 lack information on several relevant variables, including educational attainment, so we do not include them in the sample.

[13]In this calculation, we rely on reported region of residence in the 1992 census and exclude Chileans living abroad and foreign nationals. Chile is administratively divided into 346 counties located in 16 different regions.

[14]Cause of death is reported using the International Statistical Classification of Diseases and Related Health Problems (ICD), versions 9 and 10.

correspond to first-time admits, 55% to readmitted patients and 8% have a missing identifier. We use this information to construct disaggregate hospitalization rates by type of patient.

We also use the hospitalization data at the individual level to study changes in the characteristics of patients, such as type of insurance, across cohorts. Additionally, we exploit the fact that the mortality files in the vital statistics use the same individual identifier as the discharge summaries to track the mortality of hospitalized patients over multiple time horizons. Naturally, we must ommit from this analysis the small set of hospitalized patients with a missing identifier. To ensure that we observe outcomes for all hospitalized patients for at least five years, we end the sample for this part of the analysis in 2012.

Finally, we use 13 waves of Chile's National Socioeconomic Survey (CASEN) conducted between 1990 and 2017 to analyze health-related outcomes that we are unable to study using the administrative sources above.[15] The CASEN is collected biennially and records information on education, health, income and labor market outcomes. We study four families of outcomes. First, health behaviors such as smoking in the past month or, in the case of women, having a Pap test done in the past three years. Second, measures of health status such as being sick in the past three months and a self-assessment of overall health. Third, measures of access to healthcare, including having seen a general practitioner or a specialist in the past three months. Fourth, type of health insurance (public or private). CASEN is a very large survey including more than 260,000 individuals from over 80,000 households in its most recent wave. This allows us to have a relatively large sample even after restricting to individuals with 4+ years of secondary reaching age 21 between 1964 and 1981. Not every question is asked in every wave, which leads to varying sample sizes.

4 Empirical Strategy

If we could connect the individuals in the 1992 census to the mortality files in the vital statistics, a natural model to estimate the relationship between college and mortality would be as follows:

$$D_{i,t} = \beta\, C_i + \delta\, X_{i,t} + \varepsilon_{i,t}, \tag{1}$$

where $D_{i,t}$ is a dummy indicating whether individual i is deceased by time t. $X_{i,t}$ is a vector of controls that potentially vary across individuals or over time (and a constant), while C_i is a dummy equal to one for individuals with any college education. This is a fixed individual characteristic assuming we only observe people with completed education. The parameter of interest is β, which captures the average difference in the mortality rate for people that attend college over the period under study. By restricting the sample to individuals with complete secondary education, we can

[15]Survey years are 1990, 1992, 1994, 1996, 1998, 2000, 2003, 2006, 2009, 2011, 2013, 2015 and 2017.

interpret β as the average difference in mortality for people with college relative to those that stop their schooling after finishing the level immediately below, arguably the relevant counterfactual.

The model in equation (1) poses two problems. First, we must overcome the fact that we cannot link people in the census and the mortality files. Secondly, even if we could, concerns about omitted variable bias (OVB) prevent us from interpreting the OLS estimate of β as capturing the causal effect of college enrollment (Card, 1999). In our setting, OVB may arise because of unobservable differences in genetic characteristics, parental inputs or individual preferences (e.g. discount rates) that affect both the decision to go to college and health later in life (Fuchs, 1982). Previous work has shown, for instance, that low birth weight is associated with worse health during childhood and with reduced educational attainment (Behrman and Rosenzweig, 2004; Black et al., 2007). Case et al. (2002) further show that household income is positively correlated both with children's health and educational attainment.

To tackle the first problem, we aggregate the data into larger units that we observe in both data sources. As discussed in the previous section, we can collapse the data into cells at the cohort-gender-region level. Here, region refers to the region of residence, as the mortality files do not include information on place of birth.[16] Similarly to Lleras-Muney (2005), we can then derive the aggregate model by averaging the previous model over individuals in a given cell:

$$\bar{D}_{k,r,t} = \beta\,\bar{C}_{k,r,t} + \delta\,\bar{X}_{k,r,t} + \bar{\varepsilon}_{k,r,t} \tag{2}$$

$\bar{D}_{k,r,t}$ represents the share of people in a given cohort-region cell, denoted by k and r, that die in year t, while $\bar{C}_{k,r}$ represents the share of people per cell with any college.[17] By weighting each observation by the number of people in that cell, we obtain an estimate of β that is identical to the one provided by the individual microdata (Angrist and Pischke, 2009). Since people that are already deceased are no longer at risk of dying, we adjust the denominator in $\bar{D}_{k,r,t}$ and $\bar{C}_{k,r,t}$ based on the cumulative number of previous deaths per cell, following Clark and Royer (2013).

Still, any bias in the individual-level estimate of β will carry over to the cohort-level estimate. To tackle the second problem (i.e. identification), we leverage plausibly exogenous variation provided by the kink in college enrollment experienced by cohorts reaching college age in the years immediately after the 1973 military coup. In the spirit of a regression kink design (Card et al., 2015), our baseline reduced-form specification looks for a change in the cohort-level trend of the

[16]Collapsing by region allows us to account for spatial differences in mortality, while minimizing the measurement error caused by migration or misreporting. This error could be quite large in sparsely populated cells, which is why we do not use the more granular county (*comuna*) level. Reassuringly, our main results are almost identical if we do the analysis at the national level, where the threat posed by measurement error is even smaller.

[17]For simplicity, we are omitting the gender subindex in equation (2), but the collapsed cells are gender-specific. We estimate all of our models separately for men and women, but results are identical to running pooled regressions and allowing the region-year fixed effect included in all our estimations to be gender-specific.

mortality rate among those affected:

$$\bar{D}_{k,r,t} = \alpha_{r,t} + \pi_0\, k + \pi_1\, \mathbb{1}(k > 0) \times k + \eta_{k,r,t} \tag{3}$$

Here, we have normalized the indicator k to denote the year in which the cohort reaches age 21, as this is the average age of first-year college students in our setting (Bautista et al., 2020b). We have also re-scaled k to equal zero in 1972, so π_0 captures the cohort-level trend in the mortality rate, while π_1 captures any change in this trend (i.e. a kink) for cohorts reaching age 21 in 1973 or later ($k > 0$). As control, we include a region by year fixed effect, denoted by $\alpha_{r,t}$, which accounts for geographic differences in mortality and allows these differences to vary flexibly over time. It also accounts for common shocks or secular changes in health and allows their impact to vary flexibly across regions. $\eta_{k,r,t}$ is an error term clustered at the region-year level, but we also report p-values from the Wild cluster bootstrap procedure following Cameron et al. (2008) to account for clustering by cohort.

The identifying assumption in this reduced-form model is that in the absence of the military coup there should not be a kink in the cohort trend of the mortality rate for those reaching age 21 after 1973. The parsimonious specification focuses on a linear trend to avoid over-fitting and we provide visual evidence that it fits the data relatively well. While in a purely cross-sectional analysis this specification might cause concern regarding non-linear age effects, in our setting this threat is minimized by the fact that we observe the study cohorts repeatedly over 23 years. We can, however, further exploit our ability to observe cohorts repeatedly over time to estimate a more demanding specification that replaces the baseline cohort trend with an age fixed effect, $\gamma_{(k,t)}$. This way, we allow the mortality rate to vary flexibly throughout the life cycle and restrict the comparison to people from different cohorts at the same point in the cycle:

$$\bar{D}_{k,r,t} = \alpha_{r,t} + \gamma_{(k,t)} + \phi\, \mathbb{1}(k > 0) \times k + \nu_{k,r,t} \tag{4}$$

Based on either of the reduced-form models (with or without age fixed effects), we can now return to equation (2) and use the kink in college enrollment after 1973 as an excluded instrument. For the specification with age fixed effects, we estimate the following system of equations:

$$\bar{C}_{k,r,t} = \omega_{r,t} + \psi_{(k,t)} + \theta\, \mathbb{1}(k > 0) \times k + \mu_{k,r,t}, \tag{5}$$

$$\bar{D}_{k,r,t} = \alpha_{r,t} + +\gamma_{(k,t)} + \tilde{\beta}\, \bar{C}_{k,r} + \bar{\varepsilon}_{k,r,t} \tag{6}$$

where θ is the first-stage estimate of the kink in college enrollment for the affected cohorts and $\tilde{\beta}$ is the instrumental variables (IV) estimate of the causal effect of college enrollment on the mortality rate. Under standard assumptions, $\tilde{\beta}$ can be interpreted as a local average treatment effect (LATE)

(Angrist et al., 1996). This is the average causal effect of college entry for the set of compliers whose college enrollment was negatively affected by the military coup. In our setting, this is a population of particular interest as these are the (potential) college students that the technocrats guiding economic policy for the military regime argued were undeserving of educational subsidies.

We focus on college enrollment, rather than completion, because this is the margin that was directly affected by the dictatorship's policies. Bautista et al. (2020b) show that in all our study cohorts at least 68% of people with any college report four or more years of study (a proxy for completion), a figure comparable to those from other settings (e.g., Zimmerman, 2014).

Our IV strategy requires an additional exclusion restriction to be satisfied, namely, that the kink in the mortality rate for the affected cohorts is exclusively driven by the reduction in college enrollment. We find this assumption to be plausible insofar as any other changes brought about by the military regime should have affected contiguous cohorts of young adults in a roughly similar fashion. In this regard, our baseline sample focuses on a narrow bandwidth including only nine cohorts on each side of the kink, but as part of our robustness tests we show that our results are unaffected if we use tighter bandwidths (as little as four cohorts on each side). The fact that the drop in enrollment is entirely driven by the fall in openings (as shown above) and that there is no kink in secondary completion (as shown below), further indicates that the change in policy towards higher education implemented by the incoming military regime is driving the cross-cohort variation. Other events that may have affected cohorts differentially are likely, if anything, to play against us finding any effect. For instance, the cohorts that are worst affected by the reduction in college enrollment at the end of the 1970s are also the ones benefiting from a period of high economic growth. The same goes for secular improvements in health conditions that disproportionately benefit younger cohorts, such as increased awareness of the health risks of smoking. To further assuage concerns about violations of the exclusion restriction, we carry out tests following Conley et al. (2012), which reveal that such violations would have to be quite large (70% or more of the reduced-form effects) to make our findings insignificant.

5 Results: College Enrollment and Mortality

This section presents our main results on the causal effect of college enrollment on mortality. We begin by documenting a sharp kink in college enrollment for cohorts that reached college age shortly after the 1973 military coup (i.e. our first stage). We then provide reduced-form estimates of analogous kinks in the yearly mortality rate between 1994 and 2017 for these cohorts, as well as IV estimates using the post-1973 kink as an excluded instrument for college enrollment.

5.1 Educational attainment

Figure 2 plots raw data from the 1992 population census. Panel (a) shows the shares of men and women per cohort that report four or more years of secondary education (our proxy for secondary completion). These shares rise smoothly over time for both genders, starting at around 25% for the cohort reaching age 21 in 1960 (born in 1939) and peaking at around 47% for the cohort reaching age 21 in 1990 (born in 1969). Panel (b) shows the respective shares that report any college education, among those with 4+ years of secondary. Cohorts reaching age 21 before 1973 experienced a rising college enrollment rate, especially during the Allende government that begun in 1970. However, cohorts reaching the same age after the military coup (denoted by the red vertical line), experienced a sharp kink and a steady decline in the enrollment rate. While men and women reaching age 21 in 1972 had conditional enrollment rates of around 38%, those reaching the same age in 1980 had enrollment rates closer to 22% (42% decrease). The fact that secondary completion is increasing smoothly for these cohorts indicates that the drop in the enrollment rate is driven by fewer people entering higher education (i.e. the numerator). The dashed lines in both panels correspond to the start and end points for the study cohorts that we include in the analysis to follow, but the figure shows that the trends are unchanged in adjacent cohorts.

Table 1 presents estimates of equation (3) using different measures of educational attainment as dependent variable. Columns 1-3 show results for women, while columns 4-6 show the corresponding estimates for men. In columns 1 and 4, we use the estimating sample for the mortality analysis below (i.e. cohort-region-year panel). The dependent variable is the death-adjusted share of people per cell (out of every ten) that report any college education.[18] Consistent with Figure 2, the results in the top row show that college enrollment before 1973 was rising at a rate of 2 percentage points (pp) per cohort among women and 1.4 pp for men. The results in the bottom row quantify the kink in enrollment for the cohorts reaching the same age in the post-coup years. By adding the coefficients, we see that the net trend for both men and women in the affected cohorts is -2 pp. This is equivalent to a 7% yearly decrease relative to the sample mean for women, 6% for men. Columns 2 and 5 show that the estimates hardly change if we ignore the variation in the share with college caused by mortality during the sample period and estimate equation (3) using the cross-section from the 1992 census.[19] In columns 3 and 6, we use average years of college as dependent variable instead. Again, we find net negative trends among the affected cohorts. Relative to the respective sample means, affected women experience a 7% yearly decline in years of college, while men see a 5% drop.

Panels (a) and (b) in Figure 3 provide a visualization of our first stage results for women and

[18] We can interpret these as percentual changes if we multiply by ten. This normalization of the college enrollment rate facilitates the interpretation of the IV results on mortality for reasons that we discuss below.

[19] Appendix Table A1 shows analogous results for the 2002 census.

men, respectively. In these figures, the markers show average college enrollment per cohort. The solid lines indicate the estimated trends before and after the coup, while the dashed line denotes the counterfactual trend for the post-coup cohorts. The parsimonious linear model describes the evolution of the college enrollment rate across cohorts relatively well and captures the sharp negative kink for those reaching college age after 1973.

5.2 *Impact on Mortality*

Panel (a) in Table 2 provides reduced-form estimates of the kink in the risk-adjusted mortality rate for the affected cohorts. Columns 1-2 show results for women and columns 3-4 for men. Odd-numbered columns correspond to the specification that controls for the baseline cohort trend (i.e. equation 3), while even-numbered ones correspond to the more flexible specification with age fixed effects instead (equation 4). All regressions include region-year fixed effects. We report in parenthesis standard errors clustered at the region-year level, while the number in brackets corresponds to the p-value for the null of a zero coefficient when we cluster by cohort and implement the Wild bootstrap procedure following Cameron et al. (2008).

As expected, people in younger cohorts are less likely to die at any point in time.[20] Column 1 shows that the yearly number of deaths per 1,000 decreases at a rate of -0.61 per cohort among women that reached college age before 1973. For the affected cohorts, however, the mortality rate kinks upwards and decreases at the much smaller rate of -0.25 per cohort (-0.61+0.36). We observe a similar pattern for men in column 3. A baseline trend of -1.2 fewer deaths per 1,000 for each additional cohort before the coup, that flattens for the post-coup cohorts and becomes -0.44. Expressed as a percentage of the baseline trends, the measured mortality kinks for women and men in the affected cohorts equal 59% and 64% respectively. Panels (c) and (d) in Figure 3 illustrate these results. The parsimonious linear model provides a fairly accurate representation of the cohort-level trend in mortality before the coup and provides clear indication of an upward kink for the cohorts reaching age 21 after 1973.

Columns 2 and 4 show that the upward kink in mortality is present even if we replace the baseline trend with the much more flexible age fixed effects. In doing so, we restrict the comparison to cohorts at the same point in the life cycle (i.e, one-year age group) and effectively discard information from observations for which such a comparison is not possible (very old or very young). This is a much more demanding specification that expectedly absorbs a large share of the identifying variation. As a result, the estimates of the post-1973 kink decrease by a substantial amount, though they remain economically meaningful. The coefficient in column 2 indicates that each new post-coup female cohort experiences a 0.11 unit increase in the age-adjusted yearly mortality rate, equivalent to 2.8% of the sample mean of 3.9 deaths per 1,000. The corresponding estimate

[20] Panels (a) and (b) in Appendix Figure A1 show mortality profiles by age and sample year.

for men in column 4 similarly shows that each younger post-coup cohort has a 0.32-unit higher mortality rate (4.4% increase over sample mean of 7.1 deaths per 1,000). While the estimates for men remain extremely precise in this specification (statistically significant at 0.1% level), those for women are somewhat noisy once we account for clustering by cohort (p-value of 0.076).

Panel B shows the IV estimates of the effect of college enrollment. At the bottom of the table we report the Kleibergen-Paap F-statistics, which are in the several thousands and indicate that we have a very strong first stage relationship in all cases. We express the endogenous variable as the share with college per every ten individuals to be able to interpret the estimated coefficients as percentage point effects on the probability of dying. In this regard, the estimate from the baseline specification in column 1 indicates that college enrollment reduces the yearly probability of dying during the sample period by 0.88 percentage points (pp) for women, while the age-adjusted estimate in column 2 places this effect at -0.26 pp. In the case of men, the estimates in columns 3 and 4 point to reductions in the probability of dying of -2.4 pp and -0.9 pp for the specifications without and with age fixed effects. Focusing on the estimates with age fixed effects, our more conservative and preferred specification, we find that the reduction in mortality caused by college enrollment is equivalent to 68% and 130% of the respective female and male sample means.[21] As with the reduced-form estimates, the results for women are somewhat imprecise once we account for clustering by cohort (p-value of 0.091), while those for men are very precise throughout. Using the baseline clustering at the region-year level, we can reject at the 0.1% level that college reduces female and male mortality by the same amount.

Panel C reports the corresponding OLS estimates of the correlation between college enrollment and mortality.[22] Our IV estimates are larger than their OLS counterparts, but remain very much comparable. For our preferred specification with age fixed effects, IV is 39% larger than OLS for women and 14% larger for men, but these differences are not statistically significant at conventional levels.[23] Larger IV estimates are common in the returns-to-education literature (Card, 2001), even in very well-identified studies (Oreopoulos, 2006).[24] A common explanation is that the returns to college are particularly high for the complier population affected by the instrument at hand (Card, 1999). In our case, it seems plausible that the missing students that failed to enroll in university after 1973 had particularly high health returns to college, as they were at the bottom of the coun-

[21]Appendix Table A2 shows that one year of college reduces female mortality by 0.7 deaths per 1,000 and male mortality by 2.4 deaths (specifications with age fixed effects). These effects correspond to 18% and 34% of the respective sample means. Buckles et al. (2016) estimate for the US that one year of college reduces aggregate male mortality between 1981 and 2007 by 26 deaths per 1,000, equivalent to 19% of their sample mean. Hence, our estimated effect size on male mortality is somewhat larger. This comparison should be interpreted with caution, though, as it could be affected by differences in the setting, the composition of the sample and the methodology.

[22]Panel (c) in Appendix Figure A1 documents a positive correlational gradient in mortality along the entire education distribution during our sample period.

[23]In these tests, we treat the OLS estimates as fixed numbers (i.e. not a Hausman test).

[24]This patter is also common in studies specifically on the effects of education on health (Galama et al., 2018).

terfactual distribution of admitted students, with low scores and disproportionate representation of less affluent socioeconomic backgrounds.[25] Moreover, people in the affected cohorts could be affected not only by their individual reduction in the probability of college enrollment, but also by the fact that their peers experienced similar reductions in educational attainment. Our group-based (i.e. cohort-based) instrument is particularly well-suited to capture important spillover effects that the OLS estimate fails to incorporate (Grossman, 2006).[26]

An alternative explanation is that violations of the exclusion restriction artificially inflate the IV estimates. Our very strong first stage already suggests that such violations would have to be quite substantial to generate meaningful bias. To provide a more specific assessment of the robustness of our results to such violations, we carry out further tests following Conley et al. (2012). As the results at the bottom of Table 2 show, we find that the reduced-form effect of the excluded instrument (i.e. panel A) would have to be driven to a very large extent (70% or more) by factors other than reduced college enrollment for the 90% confidence interval of our IV estimates to include zero. In the next subsection, we discuss additional tests that help us rule out a meaningful impact of violations of the exclusion restriction.

5.3 Robustness checks

As part of our robustness tests, we verify that our results are unaffected if we do not adjust the mortality rate, the share with college, or both, based on the cumulative number of previous deaths in each cell. These results are presented in Appendix Tables A3-A5. We also verify in Appendix Table A6 that the results are unaffected if we estimate our models at the cohort-year level, ignoring the geographic variation across regions. This addresses potential concerns about measurement error caused by changes in location after 1992. Our results are also similar and point to meaningful reductions in mortality for people that go to college if we use the population counts in the 2002 census as starting point and study mortality between 2003 and 2017 (Appendix Table A7).

Appendix Figure A3 shows point estimates and 95% confidence intervals from a battery of regressions imposing various different changes on the composition of our sample. We focus here on the IV results from our preferred specification with age fixed effects (equation 6). Panel (a) plots the estimated effect of college enrollment on mortality as we change the set of cohorts in the sample. The rightmost set of estimates correspond to our baseline bandwidth of cohorts reaching age 21 between 1964 and 1981. Each set of estimates to the left corresponds to a one-cohort reduction on each side of the bandwidth. We find that the results are largely stable and remain negative and statistically significant throughout. In particular, the estimates with the tightest bandwidth (age 21

[25]For instance, our compliers may have been less likely to have a medical doctor in their families, which has been found to lead to worse health and higher mortality (Chen et al., 2019).

[26]The IV strategy will also help counteract the attenuation bias resulting from classical measurement error in the reported educational level in the death certificates.

between 1969-1976: four cohorts on either side) are essentially identical to our baseline results, though slightly less precise. This is important as a closer-together set of cohorts helps alleviate lingering concerns about violations of the exclusion restriction arising as a result of a differential impact of the military coup along margins other than educational attainment (or other events).

Panel (b) then examines changes to the ages in which we allow these cohorts to enter the estimating sample. Our baseline sample, corresponding to the rightmost estimates, includes cells with ages ranging from 34 to 74 between 1994 and 2017. As in panel (a), each set of estimates to the left corresponds to a one-year reduction on each end of the range of ages allowed in the sample. Once we restrict the age range to 43-65 or beyond, we are only including ages that we observe for cohorts on both sides of the 1973 kink (i.e. common support), as shown in Appendix Figure A2. Our most conservative sample, corresponding to the leftmost estimates, only allows into the sample cells with ages between 47 and 61 (15 years). Again, the estimates are quite robust as we increasingly tighten the inclusion criterion: college reduces mortality for both men and women.

Finally, panel (c) shows results as we continually reduce the end-year of the sample. As before, our baseline sample corresponds to the rightmost set of estimates, running from 1994 to 2017. We consider one-unit reductions to the end-year until 2000, which leaves us with seven years of mortality data. The estimated effect of college on mortality remains negative and statistically significant for both men and women in all samples. However, the effect of college on male mortality decreases to around -0.65 as we exclude years from the final decade of the sample. Further reductions to the sample simply make the estimate increasingly noisy, to the point that we cannot reject that the effect of college on male and female mortality is the same.

One important caveat to our main analysis above is that we are only able to study mortality starting in 1994, when the youngest of our study cohorts is 32 years old and the oldest is 51. Deaths taking place before that year, which may have disproportionately affected people in the older (i.e. pre-coup) cohorts could be leading to higher estimated mortality rates for the affected cohorts. To address this concern, we use information on marital status in the 1992 census to estimate our reduced-form and IV specifications with widow status as dependent variable.[27] Underlying this analysis is the idea that an upward kink in widowhood for the affected cohorts in 1992 reflects higher mortality among these cohorts in the previous years, under the assumption that people tend to marry others of a similar age. The results in Table A8 indeed provide evidence of such a kink, which is striking given that these are younger cohorts at a particularly young age. The IV estimates indicate that college enrollment reduces the probability of being a widow in 1992 by 8 pp for women and by 1.7 pp for men (almost five times the sample mean in both cases). These results imply that the estimates from our main panel analysis are a lower bound for the aggregate

[27]Since this analysis is purely cross-sectional, we must rely on the specifications without age fixed effects. Regression also includes an ever-married dummy to limit the comparison to people at risk of becoming widows.

effect of college enrollment on mortality for our study cohorts in the full period up to 2017.

5.4 Cause of death

Appendix Table A9 provide disaggregate estimates of the effect of college enrollment on mortality from various causes, based on the chapter of the ICD code reported in the death certificates. The set of causes included in the table is not exhaustive, but includes the main drivers of mortality in the sample: tumors, diseases of the circulatory, respiratory and digestive systems, and external causes (i.e. homicides, accidents). Panel A provides results for women, while panel B replicates the analysis for men. We focus here on the estimates from our preferred specification with age fixed effects.

The results for women show that the reduction in mortality caused by college enrollment is largely driven by deaths from cancer and diseases of the circulatory system (e.g. heart attacks) or the digestive system (e.g. cirrhosis), especially the first two. The effect of college enrollment on mortality from these three broad causes amounts to 85% of the aggregate effect. We find no impact on deaths from external causes or from diseases of the respiratory system. Appendix Table A10 further disaggregate cancer deaths into the most common types of tumor. College enrollment reduces female mortality from tumors of the digestive organs (39% of aggregate effect on cancer deaths), breast (28%), female genital organs (21%) and lymphoid tissue (25%). We find a positive but insignificant effect on tumors of respiratory organs. We return to these results when we study the mediating effect of health behaviors in the next section.

For men, the largest effects of college enrollment are also on deaths from tumors and diseases of the circulatory or digestive system (74% of the aggregate effect between them), but we also find evidence of significant reductions in deaths from external causes or from diseases of the respiratory system.[28] The disaggregate results for cancer show that the effect is concentrated in tumors of the digestive or respiratory organs or lymphoid tissue (82% of the aggregate effect on cancer mortality). Appendix Table A11 further shows that the effect on external causes is driven by non-traffic accidents (i.e. occupational), medical complications and other (unknown) causes.

Expressed as a percentage of the sample mean, we find that the reduction in cancer mortality caused by college enrollment amounts to 5.5% for women and 9.6% for men. These effect sizes are substantially smaller than the ones found for diseases of the circulatory system (11.6% for women and 15.3% for men respectively) or from diseases of the digestive system (12.5% and 22%). This pattern is consistent with the idea that idiosyncratic or hereditary factors beyond a person's control play a larger role in cancer mortality, while deaths from diseases of the respiratory or digestive system are more strongly affected by individual behaviors over a long period of time (e.g., diet,

[28] Buckles et al. (2016) similarly find that the effect of college on male mortality in the US is largely concentrated in deaths from cancer and heart disease. They provide no estimates for women.

exercise, alcohol consumption). These cause-specific effect sizes are systematically larger for men, indicating that the larger aggregate effect of college on male mortality is not driven by changes in the cause of death (i.e. competing risks). The difference across genders is particularly stark in the case of diseases of the respiratory system (17.8% for men vs an insignificant 3.7% effect for women). This could reflect a differential impact of higher education on occupational choice, whereby men without college are more likely to work in occupations with high risk for respiratory disease (mining, manufacturing). We further examine these possibilities in the next section.

6 Mechanisms

In this section, we present evidence on some of the mechanisms underlying the causal link between college enrollment and reduced mortality. The existing literature has identified many potential mechanisms (Cutler and Lleras-Muney, 2008, 2010). Borrowing the notation from Grossman (2006), we can classify these mechanisms as 'market' or 'nonmarket'. Market mechanisms relate to the well-documented fact that higher education leads to better jobs and higher income (Card, 1999). Higher income enables people to access better health care, whether of the preventive or curative type, while occupational choice can lead to differences in health hazards or access to health insurance. Nonmarket mechanisms include multiple factors related to preferences, beliefs, and skills that influence health-relevant behaviors and are potentially affected by education. For instance, education may affect health through changes in important characteristics of individual preferences, such as risk aversion or discount rates (e.g., Becker and Mulligan, 1997), or through exposure to different types of peers. Education could also increase knowledge on health-related matters or improve decision-making ability, thereby making the production of health more efficient (Grossman, 1972).

Without meaning to be exhaustive, we provide results from three sets of exercises (involving three different data sources) that shed some light on the relevance of some of these mechanisms in our setting. First, we estimate the mediating effect of labor market outcomes on the relationship between college and mortality, still relying on information from the census and the vital statistics. We then study hospitalizations, their characteristics, and the mortality of hospitalized patients, using administrative records from discharge summaries. Finally, we explore the impact of college on self-reported measures of health behaviors and health status in the CASEN household survey.

6.1 Labor market outcomes

To gain insight on the mediating effect of labor market outcomes on the link between college enrollment and reduced mortality, we rely on individual information recorded in both the 1992 census and the death certificates. For each individual, both sources report categorical information on labor force participation, type of occupation and type of employment.[29] Importantly, there are only minor differences in the way these variables are coded in both sources, allowing us to combine them and study mortality within more tightly-defined categories. Using information from each variable separately, we create a panel at the cohort-gender-category-year level, where categories are defined by the values of the respective variable. For labor force participation, we use three categories: in labor force, domestic duties, and other activities. For type of occupation, we use five: white-collar high-skill, white-collar low-skill, blue-collar, military, unemployed/inactive. The latter is also present for type of employment, which totals four categories including business owner, employee, and self-employed as well.[30] The category 'unemployed/inactive' is very highly correlated with 'domestic duties' as labor force status (correlation coefficient of 0.96 for men, 0.99 for women), so we can interpret the categorizations based on type of occupation and type of employment as more fine-grained versions of the one based on labor force status.

We calculate yearly risk-adjusted mortality rates and college shares for each cell. A valid concern about this approach is that the category reported for an individual at the time of death may easily differ from the one in the census several years before. For instance, individuals in the labor force in 1992 will inevitably retire at some point in the future.[31] To address this concern, we limit this part of the analysis to a much shorter time window covering only the first decade after the 1992 census (i.e. 1994-2002), in which the threat posed by such changes is arguably smaller, and we verify that our results are robust to small changes to this window. We also ignore the geographic variation across regions, as it could exacerbate the measurement error.

For the panel corresponding to each labor market outcome, we replicate the IV analysis from our preferred specification with age fixed effects (i.e. equation 6), replacing the set of region-year fixed effects with separate sets of year and category fixed effects. The latter absorb differences in mortality related to the labor market outcome in question. These are 'bad controls' as defined by Angrist and Pischke (2009), so we expect the estimated effect of college on mortality to decrease once we control for them, to the extent that this effect operates through them.[32] In this regard, Bautista et al. (2020b) show that decreased access to college for the affected cohorts negatively impacted them along all of the dimensions we study.

Table 3 shows the results. Column 1 replicates our baseline analysis at the cohort-year (i.e. national) level for women (panel A) and men (panel B). These results differ slightly from the ones in Table 2 because we are ignoring the geographical variation and, most importantly, because

[29] Type of occupation is reported using the International Standard Classification of Occupations (ISCO).

[30] Some of these categorizations involve combining sparsely-populated categories with larger ones for a given variable (e.g. studying or retired with other activities in labor force participation). Appendix Table B1 provides summary statistics from the 1992 census for the categories used in the analysis.

[31] The retirement age in Chile is 60 for women and 65 for men.

[32] See Cutler and Lleras-Muney (2008, 2010) for a similar strategy.

of the shorter sample period ending in 2002. Columns 2-4 then provide adjusted estimates for each panel with the corresponding set of category fixed effects.[33] We abstain from using the more stringent category by year fixed effects to make the results as comparable as possible to the baseline estimates in column 1. For the same reason, we use robust standard errors throughout.

In panel A, Columns 2 and 4 show that controlling for measures of labor force participation or type of employment reduces the magnitude of the effect of college on female mortality by 29% and 14% respectively. However, column 3 shows that controlling for type of occupation leads, in fact, to a 4% increase in the size of the college effect. This is striking as Bautista et al. (2020b) show that college enrollment has a very large effect on women's probability of having a high status occupation (i.e. white-collar high-skill). Our results indicate that the mortality college premium is smaller among women in the labor force (column 2), but suggest that it is in fact larger among women with these more prestigious occupations (column 4). These women may access preventive care less frequently due to time constraints, face higher work-related stress or have children at a later age (a known risk factor for breast cancer). More generally, these findings are consistent with previous literature on the trade-off that college-educated women often face between career success and overall well-being (Bertrand, 2013). In the case of men, columns 2-4 in panel B show that the different labor-market-category fixed effects lead to highly homogeneous decreases in the magnitude of the effect of college on mortality, ranging from 28% (column 4) to 32% (column 2).

Overall, the results in Table 3 show that labor market outcomes can explain (in a statistical sense) about 30% of the estimated causal effect of college enrollment on mortality.[34] Labor force participation seems to have the highest explanatory power, especially for women, which is consistent with both market (e.g. income and access to private health insurance) and nonmarket mechanisms (e.g. peer effects at work or through the marriage market). Further controlling for other labor market outcomes (type of occupation and employment) does not appear to add much explanatory power, which is roughly in line with previous results by Cutler and Lleras-Muney (2008, 2010). Unfortunately, data on income is unavailable in the mortality files and the census, preventing us from disentangling its mediating effect. However, we can do a back-of-the-envelope calculation based on the following system of equations:

$$\begin{aligned} & \text{Pr(Death)} = \rho \log(\text{income}) + \epsilon \\ & \text{Log(income)} = \gamma \text{ Any College} + \mu \\ \Rightarrow \quad & \text{Pr(Death)} = \underbrace{\rho \times \gamma}_{\beta} \text{ Any College} + \varepsilon \end{aligned}$$

[33] Differences in the number of observations are caused by variation in the number of categories (cells) across the three variables considered. As in our main analysis, all regressions are weighted by cell size and yield the same baseline estimates (i.e. column 1) in the absence of the category fixed effects, irrespective of the unit of observation.

[34] Appendix Figure B1 shows that the results are very similar for any end-year between 2000 and 2004.

Bautista et al. (2020b) estimate that the effect of college enrollment on log income for the affected cohorts is $\hat{\gamma} = 0.23$. Pooling across genders, we find here that $\hat{\beta} = 0.006$ (unreported estimates). If college enrollment had no effect on mortality other than through income, these estimates would imply $\rho = 0.026$ (i.e. increasing income by 10% reduces mortality by 0.26 pp). This is an implausibly large effect, equivalent to 36% of the sample mean (i.e. base probability of dying) of 0.72 pp. Moreover, the fact that occupation and type of employment add little explanatory power further suggests that the relationship between college education and reduced mortality is not exclusively driven by income.

6.2 Hospitalizations

In this section, we present results from two exercises using additional information from hospital discharge summaries. Unfortunately, this source does not contain information on educational attainment, so the analysis is exclusively reduced form. We first study hospitalizations as an intermediate health outcome that can help us understand the mediating role of consumption of health services in the documented link between reduced college enrollment and higher mortality among the cohorts affected by the military coup. In particular, failing to find evidence of a higher hospitalization rate among these cohorts, despite their worse health (i.e. higher mortality), could indicate the presence of substantial barriers in access to care. We analyze the characteristics of hospitalized patients, particularly their type of insurance, to further explore such possibilities. Secondly, we compare the mortality of hospitalized patients across the study cohorts to gain more insight on the impact of education after what could represent a substantial deterioration of health. *Ex-ante*, it is unclear whether education affects mortality among hospitalized patients.

6.2.1 Hospitalization rates

We study hospitalization rates at the cohort-gender-region-year level, as in our main analysis. Since the hospital discharge summaries are only available since 2002, we use the census from that year as the source for the initial population count per cell, which we adjust for mortality over time.

Column 1 in Table 4 shows estimates of equation (3) using the number of hospitalizations per 1,000 as dependent variable. The estimates for both genders indicate a falling hospitalization rate for younger cohorts, which kinks upwards and falls at a lower rate for those reaching college age after the military coup. Column 2 shows results from equation (4), our preferred specification with age fixed effects. Once we flexibly account for age effects, the estimated kinks become substantially smaller and are no longer statistically significant. This suggests that the linear approximation in the baseline specification may not adequately describe the evolution of hospitalizations across cohorts. Panels (a) and (e) in Figure 4 show that the cross-cohort trend in hospitalizations is noisier than the one in mortality, especially for women. There is little visual evidence of kinks in the hospitalization rate for post-coup cohorts in these plots.

The previous outcome corresponds to the number of *hospitalizations* per 1,000 individuals, irrespective of whether these represent different hospitalized patients or multiple hospital stays by the same person. Using the individual identifier in the discharge summaries, we can distinguish between the first admission per patient (per year) and subsequent readmissions during the same year. Columns 3-4 in Table 4 show results using the yearly number of *hospitalized individuals* per 1,000 as dependent variable. We find a similar pattern to the one for all hospitalizations. Once we account for age effects in column 4, there is no evidence of a kink in the hospital admission rate.[35] The raw plots in Figure 4, panels (b) and (f), again show little visual evidence of a kink.

This null result for hospitalizations is striking given the upward kink in mortality observed for the affected cohorts.[36] However, it is consistent with the idea that lower educational attainment leads these cohorts to simultaneously be in worse health and consume less medical services. These two forces push the hospitalization rate in opposite directions and roughly cancel each other out. Lower consumption of medical services could be caused by lower demand (i.e. preferences over health, discount rates) or by the existence of barriers in access to care for these individuals.

One possibility is that the reduction in educational attainment for the post-coup cohorts makes them increasingly reliant on the public health insurance (FONASA) and public hospitals, which are arguably more congested and face greater pressure on resources. To explore this possibility, we disaggregate the number of admitted patients based on their type of insurance. Columns 5-6 use the number of admissions using FONASA (per 1,000) as dependent variable, while columns 7-8 use the admission rate using private insurance (ISAPRE). For both women and men we find evidence of an upward kink in the hospitalization rate using public insurance and a comparable downward kink in the hospitalization rate using private insurance. These results are robust to the inclusion of age fixed effects in columns 6 and 8 and can be seen in panels (c), (d), (g) and (h) of Figure 4.[37] Appendix Table C5 shows similar patterns for public vs private hospitals.

[35]First admissions per patient-year represent 67% of total admissions, while readmissions within the same year correspond to 25% of the total. The remaining 8% are admissions without an individual ID. Appendix Table C1 shows that the estimated kink for the readmission rate is sensitive to the specification (if anything, negative, with age fixed effects), but there is a robust upward kink in the rate of unidentified admissions. We interpret the latter as further evidence of lower socioeconomic status and limited access to care in the affected cohorts.

[36]Appendix Table C2 replicates the main analysis for this sample (i.e. 2002 census as baseline, no restriction on education). The slight difference in sample size relative to Table 4 is caused by the absence of mortality data for 2018. Appendix Table C3 provides results on hospitalizations by cause. Focusing on the leading causes of death, the estimates with age fixed effects indicate negative kinks in cancer-related hospitalizations for both genders and no difference in the hospitalization rate for diseases of the circulatory system for men.

[37]Appendix Table C4 shows that there are also negative kinks for both genders in the rate of hospitalizations using other types of insurance (i.e. military, other private), especially for men. For women, we find some evidence of upward kinks in hospitalizations without insurance or lacking insurance information, which we take as further evidence of lower socioeconomic status and limited access to care.

These results indicate that the affected cohorts are substituting private with public healthcare, with a roughly zero net effect on the aggregate hospitalization rate. Moreover, they suggest that differences in access to care or in the quality of services received, related to increased reliance on the more congested and less generous public health system, contribute to greater mortality in these cohorts. Though we cannot rule out that higher reliance on public health also reflects differences in preferences (i.e. FONASA as a voluntary rather than 'forced' choice by less-educated individuals), it seems more plausible that the affected cohorts are being priced out of the ever more expensive private health system, as discussed in section 2. To further understand the contribution of these factors to our mortality results, we turn next to the mortality of hospitalized patients.

6.2.2 Mortality of Hospitalized patients

To be able to exploit the availability of individual-level covariates in the discharge summaries, we study the mortality of hospitalized patients at the individual level. We drop hospitalizations lacking a patient ID (8% of total) and restrict the sample to the first observation per patient in the hospital discharge summaries (though we do take readmissions into account below). We merge these records with the mortality files and establish for each individual whether death occurs over various time horizons, ranging from time of discharge to five years. To ensure that we observe deaths for all patients over all the possible time horizons, we end the sample of hospitalizations in 2012 and track the mortality of these patients until 2017, the last year with mortality data. By looking at different time horizons, we can study whether any observed changes in the mortality of hospitalized patients in the affected cohorts correspond to temporary effects that wash away over longer time periods (i.e. dying within a year rather than within the next five).[38] We estimate a modified version of our reduced-form models that replaces the region-year fixed effects with more conservative county-year fixed effects and sequentially add additional controls to measure the mediating effect of various observable characteristics from the hospitalization records.[39] We cluster standard errors two-way by county and region-year following Cameron et al. (2011).

Table 5 presents the results for deaths within one year of hospitalization. We focus here on the one-year horizon to maximize comparability with our estimates of yearly mortality in the main analysis, but Appendix Tables C6-C10 show equivalent estimations for other time horizons. The average one-year mortality rate in this sample is 46 deaths per 1,000 admitted women and 78 deaths per 1,000 admitted men. These rates are one order of magnitude larger than the averages from our main sample in Table 2 and indicate, perhaps unsurprisingly, that hospitalized patients are at a much higher risk of death than the population at large.[40]

[38] Alas, all changes in mortality are temporary over a sufficiently long period of time.

[39] Information on some characteristics in the hospital discharge summaries is unavailable before 2004, so we set this as the initial year of the sample for this part of the analysis.

[40] Appendix Table C11 replicates our main analysis for the period 2004-2012 to further enhance comparability.

Column 1 shows estimates of equation (3) and provides evidence of an upward kink in mortality for women (panel A) and men (panel B) in the affected cohorts. These kinks are quite sizable, at 49% and 45% of the respective female and male baseline trends. Figure 5 plots the raw data and provides compelling visual evidence of the upward shift in the mortality trend.[41] Column 2 corresponds to equation (4), our preferred specification with age fixed effects. The estimated kinks decrease slightly once we account for age effects, but remain largely comparable to the baseline results. We take these estimates as the benchmark as we proceed to examine the mediating effect of various characteristics on the higher mortality of hospitalized patients in the affected cohorts.

We first examine whether differences in the cause of hospitalization help explain the kink in mortality. Column 3 includes a full set of diagnostic fixed effects, using the 4-digit ICD code from the discharge summaries. Columns 4-8 then sequentially include additional sets of fixed effects for other observable characteristics that may also be part of the mechanism. In column 4, we account for the type of insurance (FONASA or not), while in column 5 we control for the specific hospital of admission (381 different establishments) and in column 6 for the type of admission (emergency room, transfer from other hospital, etc.). Column 7 further controls for whether surgery was performed, while column 8 accounts for whether the patient had been previously admitted before the start of the sample period (i.e. 2002-2003). To maximize explanatory power, we estimate saturated models that interact each additional set of fixed effects with all the previous ones.[42] This means that in column 5, for instance, we are comparing patients across cohorts that share the same diagnostic and were admitted to the same hospital with the same insurance, and so on.

A comparison of columns 2 and 8 reveals that these observable characteristics together explain (in a statistical sense) 36% of the upward kink in one-year mortality for women in the affected cohorts and 41% for men. Among these characteristics, just accounting for differences in diagnostic makes the magnitude of the estimated kink drop by about 24% for both men and women. This suggests that different morbidity profiles (i.e. differences in the incidence of cancer or diseases of the circulatory system) are more important than the type or quality of medical services received once admitted to hospital. The remaining roughly 60% of the excess mortality that we cannot account for is arguably explained by unobservable factors, such as health behaviors and access to preventive (i.e. primary) care, which could also contribute to differences in the incidence of disease (e.g. Pap smear and cervical cancer in women). Some of the unexplained variation could also correspond to the inevitable coarseness of some of our controls. For example, diagnostic codes may not account for nuanced differences in diagnostic that affect prognosis.

Figure 6 plots the estimated kink from the regressions in Table 5 and the corresponding esti-

[41] Appendix Figure C1 shows the corresponding raw plots for other time horizons.

[42] While we keep the sample unchanged in all columns, our estimation package drops singleton observations, which become more frequent as we include increasingly stringent combinations of fixed effects (Correia, 2015). Naturally, the point estimates are unaffected by this exclusion.

mates as we change the time horizon over which we study mortality. Except for the estimates for the one-month mortality rate for women, almost all other coefficients are statistically significant at the 5% level (see Appendix Tables C6-C10). For both genders, we find larger kinks over longer time horizons, indicating that the effect of education on the mortality of hospitalized patients persists and grows for at least five years, rather than simply shifting the time of death in the short run. Despite what appear to be steeper lines (i.e larger reductions in the kink) over longer time horizons, especially for men, the explanatory power of the observable characteristics we consider is largely stable. For example, observables explain 42% of the kink in five-year mortality for women and 47% of the kink for men.[43] Differences in diagnostic explain 20-25% of the kink throughout.

6.3 Survey responses

In this section, we present a final set of results based on individual responses to the CASEN household survey between 1990 and 2017. The CASEN survey allows us to gain insight on the consumption of medical services other than hospital admissions, as well as on health behaviors and outcomes other than mortality. The information in CASEN is all self-reported, which raises obvious concerns about biases in reporting, though we have no reason to expect differential reporting across our study cohorts. Moreover, not much is known about whether such biases vary with education (Cutler and Lleras-Muney, 2010). CASEN records information on educational attainment, enabling us to restrict the sample to individuals with four or more years of secondary education and to provide both reduced form and IV results. To maximize the precision of our estimates, we pool men and women in our main analysis and reserve disaggregate results by gender for the online appendix. Similarly to the previous analysis of the mortality of hospitalized patients, we estimate our models at the individual level and include county by year by gender fixed effects in all regressions, which account for a host of spatial and temporal factors. We cluster standard errors two-way by county and region-year following Cameron et al. (2011).

Table 6 shows results on health insurance and consumption of medical services. Odd-numbered columns are based on the specification with a linear cohort trend, while even-numbered columns present results from our preferred specification with age fixed effects. The dependent variable in columns 1-2 is a dummy equal to one if the individual reports being enrolled in the public health insurance (FONASA). Information on health insurance has been asked in every wave of CASEN, leading to a sample of over 160,000 individuals. The reduced-form estimates in panel A show that there is a downward cohort trend in FONASA enrollment (i.e. younger cohorts less likely to be enrolled) that fully reverses and becomes positive for the affected cohorts (column 1). This kink is almost unaffected by the introduction of age fixed effects in column 2 and is clearly visible in the raw data shown in panel (a) of Figure 7. The corresponding IV estimate in panel B shows

[43] In the short run (e.g., one month), the combined explanatory power of the observables exceeds 50%.

that college enrollment reduces FONASA enrollment by 30 percentage points (pp), equivalent to 50% of the sample mean, which is only slightly larger than its OLS counterpart in panel C. This result is consistent with the upward kink in hospitalizations using FONASA documented above and indicates that university enrollment has a very large effect on access to the private health system.[44]

The dependent variables in columns 3-8 are respective dummies for visits in the past three months to a general practitioner (GP), the Emergency Room (ER), or a specialist physician.[45] The raw data in panels (b)-(d) of Figure 7 provides weak evidence of kinks, especially for ER visits. As with hospitalizations above, this is consistent with the affected cohorts having worse health but also more limited access to health services, which largely offset each other. Once we account for age effects in columns 4, 6 and 8, we find a null effect on ER visits together with evidence of stronger negative kinks in visits to the GP or the specialist. The IV estimates indicate that college enrollment raises the probability of seeing a GP in the past three months by 10.6 pp and of seeing a specialist by 7.7 pp. These results are consistent with more educated individuals making larger investments in health capital to increase their time available for market and non-market activities (Grossman, 1972), and with previous research documenting a positive correlation between education and the use of preventive care services (e.g., Fletcher and Frisvold, 2009; Cutler and Lleras-Muney, 2010; Lange, 2011). The null result for ER visits is consistent with the existence of fewer barriers to access of this service compared to outpatient consultations, as well as with reduced heterogeneity in the demand for it (i.e. a visit to the ER is often unavoidable given a large negative health shock). Moreover, less-educated individuals may be partially substituting outpatient consultations with ER visits, thereby widening the gap in the estimated effect across services.[46]

Table 7 has a similar structure to Table 6, but shows results on health behaviors and self-reported health status.[47] The dependent variable in columns 1-2 is a dummy equal to one if the individual reports smoking in the past month. Despite the small sample size (smoking was only asked in the 1990, 1992 and 1998 waves of CASEN), the reduced-form estimates in panel A show that the positive cohort trend in smoking (i.e. younger cohorts smoking more) disappears almost entirely for the cohorts reaching college age after 1973. This downward kink in smoking is also robust to the inclusion of age fixed effects and can be easily seen in panel (e) of Figure 7. The IV estimate in panel B points to a 41 pp positive effect of college enrollment on the probability of

[44] Appendix Tables D1-C2 show that the effect of college on FONASA enrollment is twice as large for men than for women, plausibly because women can more easily access private health insurance through their partner or spouse. This is one potential channel that could explain the larger effect of college on male mortality documented above.

[45] Sample size varies due to changes in the questions on consumption of health services across survey waves. Appendix Table C3 shows results for other health services, which are largely similar to the ones we present here.

[46] Appendix Tables D1-C2 show that the effect of college on GP visits is entirely driven by men, while the effect on specialist visits is entirely driven by women. This suggests that different mechanisms could be driving the link between college and mortality across genders. i.e. women without college are not less likely access primary care, but struggle to access a specialist, which may lead to increased mortality from conditions such as cancer.

[47] Appendix Tables C4-C5 provide disaggregate results by gender.

smoking, which goes in the opposite direction of the weakly negative OLS estimate in panel C.

This result is striking both because of its sign and its magnitude. A substantial literature has documented a negative correlation between education and smoking, but whether this corresponds to a causal effect remains unclear. De Walque (2007) and Grimard and Parent (2007) provide evidence of a negative causal effect, but both studies are based in the same setting and rely on the same source of variation, the avoidance of the military draft by young males in the US during the Vietnam war. Contextual differences could partly explain the heterogeneous impact of college across settings, given that while only 21% of Americans smoked in 2005, 44% of Chileans did in 2003, which suggests large differences in the stigma associated with smoking.[48] Additionally, other findings in the literature suggest that the negative correlation between smoking and education may be uninformative as to the causal effect.[49] The positive LATE in our setting could reflect that enrollment in university exposed individuals to a more liberal or bohemian culture, in which behaviors such as smoking were more tolerated or even fostered.[50] This is consistent with evidence by Cutler and Lleras-Muney (2008, 2010) that more educated people are more likely to drink alcohol or to have ever tried marijuana. Moreover, the large difference between the IV and OLS estimates could reflect the fact that our instrument for reduced college enrollment is group- (i.e. cohort-) based and that smoking is very strongly affected by the behavior of peers (Gaviria and Raphael, 2001; Nakajima, 2007). The significance of this finding for our main results is that the higher mortality in the affected cohorts occurs despite their sharp reduction in smoking, which is one of the main contributors to premature adult mortality. This suggests that reduced college enrollment has large effects through other channels, such as reduced access to health services.

The sample in columns 3-4 is restricted to women and the dependent variable is a dummy for having had a Pap smear (the main procedure to test for cervical cancer) in the past three years. Panel (f) in Figure 7 provides clear evidence of a downward kink for the affected cohorts, which the regression estimates show is substantial and robust to the inclusion of age fixed effects (column 4). We estimate in panel B that college enrollment increases the probability of having had this procedure by 16.8 pp, equivalent to 23% of the sample mean. This arguably contributes to the negative impact of college on deaths from female genital cancer documented in Appendix Table A10. CASEN also asks women that report not having had a Pap smear the reason for not doing

[48] Figures from the Center for Disease control (CDC) for the US and from the Ministry of Health in Chile. Recent waves of the Chilean National Health Survey (*Encuesta Nacional de Salud, ENS*) show in fact a positive education gradient in smoking (MS, 2017). Unfortunately, the sample size of this survey is too small for our analysis.

[49] Farrell and Fuchs (1982) show that differences in smoking related to years of higher education are present before the schooling is realized. Park and Kang (2008) find no effect of education on smoking in South Korea. The theoretical model in Galama et al. (2018) further suggests that diminishing returns to wealth may cause education to raise the demand for unhealthy consumption.

[50] Bautista et al. (2020b) show that around two thirds of students were enrolled in public (non-religious) universities at the time of the coup and that these universities were the ones more strongly affected by the reduction in openings.

so, which we use as additional dependent variables in Appendix Table C6. We find that negligence/forgetfulness explains 45% of the college effect, while lack of knowledge about the test or lack of interest in getting it explain a further 18%. These results should be interpreted with caution, as they are imprecise and the stated cause could be affected by framing (i.e. negligence could be masking differences in barriers to access). However, they suggest that college does impact health through changes in relevant knowledge and preferences.

Back in Table 7, columns 5-6 study whether the individual reports being in good health, while columns 7-8 look at whether the individual reports having been sick or in an accident in the past three months. Panel (g) in Figure 7 does suggest the presence of an upward kink in self-reported good health, but this appears to be driven by the cohorts at the extremes and the results in column 7 show no significant effect once we restrict the comparison to people of the same age. The fact that there is no robust change in self-reported health despite the documented upward kink in mortality among the affected cohorts suggests that health self-assessments may be unreliable, especially when not attached to specific measures of morbidity or human function (Clarke and Ryan, 2006). Columns 7-8 also show no meaningful impact on the probability of being recently sick or in an accident. This last result is consistent with our previous finding of no effect on the probability of a visit to the emergency room.

7 Concluding Remarks

In this paper, we exploit the sharp negative kink in college enrollment experienced by Chilean cohorts reaching college age shortly after the 1973 military coup to provide novel evidence on the impact of higher education on mortality. We find that women and men in these cohorts have higher age-adjusted mortality rates between 1994 and 2017, mostly related to diseases of the circulatory system and cancer. Labor market outcomes explain about 30% of these effects. The excess mortality in the affected cohorts is also observed among hospitalized patients over multiple time horizons, with observable characteristics such as diagnostic and type of insurance or hospital explaining about 40%. Differences in the consumption of health services related to the type of insurance appear to play an important role. We find mixed evidence on health behaviors, with college increasing the probability of smoking while also increasing preventive behaviors.

These findings have important policy implications. In particular, they indicate that conventional estimates of the return to college, based exclusively on the pecuniary returns, may be underestimating the aggregate return by ignoring the effects of education on nonmarket outcomes such as health (Rodríguez et al., 2015). Upwards adjustments can inform the debate on the desirability of government intervention and funding of higher education, especially since individuals may fail to internalize nonpecuniary returns.

References

Albouy, V. and Lequien, L. (2009). Does Compulsory Education Lower Mortality? *Journal of Health Economics*, 28(1):155 – 168.

Angrist, J. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics*. Princeton University Press.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of Causal Effects using Instrumental Variables. *Journal of the American Statistical Association*, 91(434):444–455.

Arendt, J. N. (2005). Does Education Cause Better Health? A Panel Data Analysis Using School Reforms for Identification. *Economics of Education Review*, 24(2):149–160.

Bautista, M. A., González, F., Martínez, L. R., Muñoz, P., and Prem, M. (2020a). The Geography of Repression and Opposition to Autocracy. *American Journal of Political Science*.

Bautista, M. A., González, F., Martínez, L. R., Prem, M., and Muñoz, P. (2020b). Technocratic Autocracy, Higher Education and Social Mobility. Working Paper.

Becker, G. S. and Mulligan, C. B. (1997). The Endogenous Determination of Time Preference. *Quarterly Journal of Economics*, 112(3):729–758.

Behrman, J. R. and Rosenzweig, M. R. (2004). Returns to Birthweight. *Review of Economics and Statistics*, 86(2):586–601.

Beltrán-Sánchez, H., Finch, C. E., and Crimmins, E. M. (2015). Twentieth Century Surge of Excess Adult Male Mortality. *Proceedings of the National Academy of Sciences*, 112(29):8993–8998.

Bertrand, M. (2013). Career, Family, and the Well-Being of College-Educated Women. *American Economic Review*, 103(3):244–50.

Black, D. A., Hsu, Y.-C., and Taylor, L. J. (2015). The Effect of Early-life Education on Later-life Mortality. *Journal of Health Economics*, 44:1 – 9.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2007). From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes. *Quarterly Journal of Economics*, 122(1):409–439.

Braakmann, N. (2011). The Causal Relationship between Education, Health and Health Related Behaviour: Evidence from a Natural Experiment in England. *Journal of Health Economics*, 30(4):753–763.

Brunner, J. J. (1984). Informe Sobre el Desarrollo y el Estado Actual del Sistema Universitario en Chile. Programa Flacso-Santiago de Chile, Documento de Trabajo.

Buckles, K., Hagemann, A., Malamud, O., Morrill, M., and Wozniak, A. (2016). The Effect of College Education on Mortality. *Journal of Health Economics*, 50:99 – 114.

Cameron, C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics*, 90(3):414–427.

Cameron, C., Gelbach, J. B., and Miller, D. L. (2011). Robust Inference with Multi-way Clustering. *Journal of Business and Economic Statistics*, 29(2):238–249.

Card, D. (1999). The Causal Effect of Education on Earnings. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 3, pages 1801 – 1863. Elsevier.

Card, D. (2001). Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica*, 69(5):1127–1160.

Card, D., Lee, D. S., Pei, Z., and Weber, A. (2015). Inference on Causal Effects in a Generalized Regression Kink Design. *Econometrica*, 83(6):2453–2483.

Card, D. and Lemieux, T. (2001). Going to College to Avoid the Draft: The Unintended Legacy of the Vietnam War. *American Economic Review*, 91(2):97–102.

Case, A., Lubotsky, D., and Paxson, C. (2002). Economic Status and Health in Childhood: The Origins of the Gradient. *American Economic Review*, 92(5):1308–1334.

Chen, Y., Persson, P., and Polyakova, M. (2019). The Roots of Health Inequality and the Value of Intra-family Expertise. Technical report, National Bureau of Economic Research.

Clark, D. and Royer, H. (2013). The Effect of Education on Adult Mortality and Health: Evidence from Britain. *American Economic Review*, 103(6):2087–2120.

Clarke, P. M. and Ryan, C. (2006). Self-Reported Health: Reliability and Consequences for Health Inequality Measurement. *Health Economics*, 15(6):645–652.

Clínicas de Chile (2016). Dimensionamiento del Sector de Salud Privado de Chile: Actualización a Cifras 2016. Technical Report.

Comisión Rettig (1996). *Informe de la Comisión Nacional de Verdad y Reconciliación*. Chile: Ministerio del Interior, Corporación Nacional de Reparación y Reconciliación.

Conley, T., Hansen, C., and Rossi, P. (2012). Plausibly Exogenous. *Review of Economics and Statistics*, 94(1):260–272.

Correia, S. (2015). Singletons, Cluster-Robust Standard Errors and Fixed Effects: A Bad Mix. Working Paper.

Cutler, D. M. and Lleras-Muney, A. (2008). Education and Health: Evaluating Theories and Evidence. In Schoeni, R., House, J., Kaplan, G., and Pollack, H., editors, *Making Americans Healthier: Social and Economic Policy as Health Policy*, pages 29–60. Russell Sage Foundation.

Cutler, D. M. and Lleras-Muney, A. (2010). Understanding Differences in Health Behaviors by Education. *Journal of Health Economics*, 29(1):1–28.

De Walque, D. (2007). Does Education Affect Smoking Behaviors?: Evidence using the Vietnam Draft as an Instrument for College Education. *Journal of Health Economics*, 26(5):877–895.

Duarte, F. (2011). Switching Behavior in a Health System with Public Option. Working Paper.

Farrell, P. and Fuchs, V. R. (1982). Schooling and Health: The Cigarette Connection. *Journal of Health Economics*, 1(3):217 – 230.

Fletcher, J. M. and Frisvold, D. E. (2009). Higher Education and Health Investments: Does More Schooling Affect Preventive Health Care Use? *Journal of Human Capital*, 3(2):144–176.

Fuchs, V. R. (1982). Time Preference and Health: An Exploratory Study. In Fuchs, V. R., editor, *Economic Aspects of Health*, pages 93–120. University of Chicago Press.

Galama, T., Lleras-Muney, A., and van Kippersluis, H. (2018). The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-experimental Evidence. Oxford Research Encyclopedia of Economics and Finance.

Galetovic, A. and Sanhueza, R. (2013). Un Análisis Económico de la Integración Vertical entre Isapres y Prestadores. Working Paper.

Gaviria, A. and Raphael, S. (2001). School-Based Peer Effects and Juvenile Behavior. *Review of Economics and Statistics*, 83(2):257–268.

Grimard, F. and Parent, D. (2007). Education and Smoking: Were Vietnam War Draft Avoiders Also More Likely to Avoid Smoking? *Journal of Health Economics*, 26(5):896–926.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 80(2):223–255.

Grossman, M. (2006). Education and Nonmarket Outcomes. In Hanushek, E. A. and Welch, F., editors, *Handbook of the Economics of Education*, volume 1, pages 577–633. Elsevier.

Kemptner, D., Jürges, H., and Reinhold, S. (2011). Changes in Compulsory Schooling and the Causal Effect of Education on Health: Evidence from Germany. *Journal of Health Economics*, 30(2):340–354.

Kitagawa, E. M. and Hauser, P. M. (1968). Education Differentials in Mortality by Cause of Death: United States, 1960. *Demography*, 5(1):318–353.

Lange, F. (2011). The Role of Education in Complex Health Decisions: Evidence from Cancer Screening. *Journal of Health Economics*, 30(1):43–54.

Levy, D. (1986). Chilean Universities under the Junta: Regime and Policy. *Latin American Research Review*, 21(3):95–128.

Lleras-Muney, A. (2005). The Relationship Between Education and Adult Mortality in the United States. *Review of Economic Studies*, 72(1):189–221.

Mazumder, B. (2008). Does Education Improve Health? A Reexamination of the Evidence from Compulsory Schooling Laws. *Economic Perspectives*, 32(2).

MDS (2018). Síntesis de resultados Encuesta CASEN 2017: Salud. Ministerio de Desarrollo Social.

Meghir, C., Palme, M., and Simeonova, E. (2018). Education and Mortality: Evidence from a Social Experiment. *American Economic Journal: Applied Economics*, 10(2):234–56.

Montez, J. K., Hummer, R. A., and Hayward, M. D. (2012). Educational Attainment and Adult Mortality in the United States: A Systematic Analysis of Functional Form. *Demography*, 49(1):315–336.

MS (2017). Encuesta Nacional de Salud: 2016-2017. Ministerio de Salud.

Nakajima, R. (2007). Measuring Peer Effects on Youth Smoking Behaviour. *Review of Economic Studies*, 74(3):897–935.

OECD (2019). Health at a Glance 2019: OECD indicators. *OECD Publishing*.

Oreopoulos, P. (2006). Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter. *American Economic Review*, 96(1):152–175.

Oreopoulos, P. and Salvanes, K. G. (2011). Priceless: The Nonpecuniary Benefits of Schooling. *Journal of Economic Perspectives*, 25(1):159–84.

Pardo, C. and Schott, W. (2012). Public Versus Private: Evidence on Health Insurance Selection. *International Journal of Health Care Finance and Economics*, 12:39–61.

Park, C. and Kang, C. (2008). Does Education Induce Healthy Lifestyle? *Journal of Health Economics*, 27(6):1516 – 1531.

PIIE (1984). *Las Transformaciones Educacionales Bajo el Régimen Militar, Vols. I y II*. Programa Interdisciplinario de Investigaciones en Educación.

Rodríguez, J., Urzúa, S., and Reyes, L. (2015). Heterogeneous Economic Returns to Post-Secondary Degrees: Evidence from Chile. *Journal of Human Resources*, 51(2):416 – 460.

Universidad de Chile (1972). *La Universidad de Chile: Antecedentes e Informaciones*. Editorial Universitaria S.A.

Universidad de Chile (2011). Compendio Estadístico - Proceso de Admisión Año Académico 2011. Vicerectoría de Asuntos Académicos.

Valdés, J. G. (1995). *Pinochet's Economists: The Chicago School in Chile*. Cambridge University Press.

Van Kippersluis, H., O'Donnell, O., and Van Doorslaer, E. (2011). Long-run Returns to Education: Does Schooling Lead to an Extended Old Age? *Journal of Human Resources*, 46(4):695–721.

Zimmerman, S. D. (2014). The Returns to College Admission for Academically Marginal Students. *Journal of Labor Economics*, 32(4):711–754.

Table 1: Educational Attainment

	Female			Male		
Dependent variable:	Share w/ college per 10 people	Share w/ college per 10 people	Average years of college	Share w/ college per 10 people	Share w/ college per 10 people	Average years of college
	(1)	(2)	(3)	(4)	(5)	(6)
Yr Age 21	0.203***	0.201***	0.080***	0.143***	0.150***	0.055***
	(0.004)	(0.013)	(0.005)	(0.002)	(0.008)	(0.004)
Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)	-0.405***	-0.398***	-0.156***	-0.332***	-0.330***	-0.130***
	(0.006)	(0.019)	(0.007)	(0.003)	(0.013)	(0.006)
Year x Region FE	Yes	No	No	Yes	No	No
Region FE	No	Yes	Yes	No	Yes	Yes
Adjusted for deaths	Yes	No	No	Yes	No	No
Observations	6,480	270	270	6,480	270	270
R-squared	0.894	0.895	0.896	0.911	0.917	0.909
Mean DV	2.783	2.759	1.167	3.207	3.150	1.397

Notes: The unit of analysis is cohort-region-year in columns 1 and 4, and cohort-region in columns 2-3, 5-6. Observations are weighted by cell size. Dependent variable in the header. Original sample includes all respondents of the 1992 census from cohorts born between 1943 and 1960 (both inclusive) that report full secondary education. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached 21 years of age, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. As additional controls, columns 1 and 4 include region by year fixed effects, while columns 2-3 and 5-6 include region fixed effects. In columns 1 and 4, the share with college is adjusted to reflect previous mortality by educational attainment. Standard errors clustered by region-year in parentheses in columns 1 and 4, robust standard errors otherwise. *** p<0.01, ** p<0.05, * p<0.1

Table 2: Mortality Rate

	Dependent variable: Deaths per 1,000 people			
	Female		Male	
	Trend	Age FE	Trend	Age FE
	(1)	(2)	(3)	(4)
	Panel A: Reduced Form			
Yr Age 21	-0.610***		-1.244***	
	(0.031)		(0.064)	
	[0.000]		[0.000]	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	0.357***	0.107***	0.800***	0.315***
	(0.024)	(0.022)	(0.046)	(0.042)
	[0.000]	[0.076]	[0.000]	[0.000]
	Panel B: IV			
Share with college per 10 people	-0.880***	-0.263***	-2.409***	-0.922***
	(0.057)	(0.054)	(0.137)	(0.121)
	[0.001]	[0.091]	[0.001]	[0.001]
	Panel C: OLS			
Share with college per 10 people	-0.681***	-0.189***	-1.662***	-0.810***
	(0.045)	(0.046)	(0.098)	(0.102)
	[0.000]	[0.100]	[0.001]	[0.001]
Year x Region FE	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes
Observations	6,480	6,480	6,480	6,480
R-squared (panel A)	0.677	0.756	0.749	0.850
R-squared (panel C)	0.672	0.755	0.738	0.850
Mean DV	3.850	3.850	7.084	7.084
Kleibergen-Paap F-statistic (panel B)	5015	4574	9875	8783
Exclusion restriction test (% of RF)	88.2%	70.1%	88.8%	85%
H_0 : OLS = IV (p-value)	0.001	0.172	0.000	0.355

Notes: Unit of analysis is cohort-region-year. Observations weighted by cell size. Sample includes all respondents in the 1992 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) reported full secondary education. Sample period: 1994-2017. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. Mortality rate and share with college adjusted for previous mortality. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1. P-value from Wild cluster bootstrap by cohort reported in brackets.

Table 3: Mediating Effect of Labor Market Outcomes

	Dependent variable: Deaths per 1,000 people			
		Fixed effects		
	Baseline	Labor force participation	Occupation category	Employment type
	(1)	(2)	(3)	(4)
		Panel A: Female		
Share with college per 10 people	-0.274***	-0.195*	-0.286**	-0.236**
	(0.071)	(0.113)	(0.142)	(0.113)
		Panel B: Male		
Share with college per 10 people	-0.685***	-0.467**	-0.497**	-0.491**
	(0.154)	(0.190)	(0.226)	(0.196)
Year FE	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes
Category FE	No	Yes	Yes	Yes
Observations	160	486	810	648
Mean DV (panel A)	1.93	1.92	1.93	1.92
Mean DV (panel B)	3.54	3.55	3.56	3.52
Kleibergen-Paap F-statistic (panel A)	1081	799	2163	768
Kleibergen-Paap F-statistic (panel B)	1579	1239	1088	1870

Notes: The unit of analysis is cohort-year in column 1 and cohort-region-category in columns 2-4. Observations are weighted by cell size. Categories determined by the labor market outcome in the header. Column 2: In labor force, domestic duties, other activities. Column 3: white-collar high-skill, white-collar low-skill, blue-collar, military, unemployed/inactive. Column 4: business owner, employee, self-employed, unemployed/inactive. Population at risk includes all respondents in the 1992 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) reported full secondary education. Sample period: 1994-2002. The interaction of "Yr Age 21", a continuous variable indicating the year at which the cohort reached age 21 (normalized to zero in 1972) with a dummy for cohorts that reached age 21 on or after 1973 is used as excluded instrument for the share with college. Mortality rate and share with college adjusted for previous mortality. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Hospitalizations

	Dependent variable: Hospital admissions per 1,000							
	All		Admitted patients					
	admissions		All		Public insurance		Private insurance	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Female				
Yr Age 21	-3.883***		-2.295***		-2.605***		0.357***	
	(0.122)		(0.094)		(0.107)		(0.025)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	1.431***	-0.097	1.108***	-0.071	1.533***	0.245**	-0.398***	-0.295***
	(0.110)	(0.133)	(0.072)	(0.089)	(0.076)	(0.101)	(0.033)	(0.040)
				Panel B: Male				
Yr Age 21	-7.056***		-4.185***		-4.142***		0.017	
	(0.152)		(0.107)		(0.134)		(0.034)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	2.283***	0.312	1.498***	0.036	2.063***	0.565***	-0.374***	-0.322***
	(0.134)	(0.193)	(0.080)	(0.114)	(0.082)	(0.145)	(0.043)	(0.050)
Year x Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4,590	4,590	4,590	4,590	4,590	4,590	4,590	4,590
R-squared (panel A)	0.822	0.857	0.856	0.894	0.869	0.919	0.909	0.919
R-squared (panel B)	0.909	0.930	0.918	0.943	0.912	0.945	0.918	0.922
Mean DV (panel A)	119	119	79.51	79.51	62.36	62.36	10.30	10.30
Mean DV (panel B)	123.5	123.5	80.21	80.21	60.17	60.17	11.20	11.20

Notes: The unit of analysis is cohort-region-year. Observations weighted by cell size. Dependent variable in the header. Number of hospitalizations (events) in columns 1-2. Number of hospitalized patients (with ID) in columns 3-4. Number of patients using public insurance (FONASA) in columns 5-6 and patients using private insurance (ISAPRE) in columns 7-8. Hospitalization rates adjusted for previous mortality within cell. Population at risk includes all respondents in the 2002 census that reached age 21 between 1964 and 1981 (both years inclusive). Sample period: 2002-2018. "Yr Age 21" is a continuous variable indicating the year when the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: One-year Mortality Rate of Hospitalized Patients (2004-2012)

	Dependent variable: Deaths within one year of discharge per 1,000 hospitalized patients							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Female				
Yr Age 21	-5.345***							
	(0.200)							
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	2.627***	1.794***	1.379***	1.313***	1.271***	1.260***	1.205***	1.145***
	(0.235)	(0.374)	(0.296)	(0.307)	(0.337)	(0.338)	(0.320)	(0.350)
				Panel B: Male				
Yr Age 21	-6.255***							
	(0.199)							
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	2.817***	2.217***	1.643***	1.488***	1.524***	1.526***	1.440***	1.297***
	(0.275)	(0.360)	(0.354)	(0.354)	(0.399)	(0.403)	(0.401)	(0.400)
Year x County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects: Diagnostic...	No	No	Yes	No	No	No	No	No
... x Public insurance...	No	No	No	Yes	No	No	No	No
... x Hospital...	No	No	No	No	Yes	No	No	No
... x Type of admission...	No	No	No	No	No	Yes	No	No
... x Surgery...	No	No	No	No	No	No	Yes	No
... x Previously admitted	No	No	No	No	No	No	No	Yes
Observations (panel A)	603,878	603,878	602,862	601,722	539,443	509,174	493,174	468,553
Observations (panel B)	519,615	519,615	518,681	517,564	455,207	426,039	410,436	390,574
R-squared (panel A)	0.019	0.019	0.237	0.247	0.344	0.374	0.387	0.401
R-squared (panel B)	0.019	0.019	0.254	0.266	0.363	0.389	0.401	0.412
Mean DV (panel A)	45.82	45.82	45.80	45.80	43.38	41.96	41.05	40.10
Mean DV (panel B)	77.82	77.82	77.82	77.82	76.20	74.89	74.29	73.30

Notes: The unit of analysis is a hospitalized patient. Dependent variable is an indicator for whether the patient dies within one year of discharge (multiplied by 1,000). Sample includes patients from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and is limited to one observation per patient (i.e. first admission). Sample period: 2004-2012. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. All regressions include year by county of residence fixed effects. Columns 2-8 replace the baseline cohort trend with age fixed effects. Columns 3-8 iteratively add interactions of fixed effects for diagnostic (column 3: 4-digit ICD code), public insurance (column 4), hospital (column 5: 381 establishments), type of admission (column 6: ER, other establishment, etc.), surgery (column 7), previously admitted (column 8: 2002-2003). Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Health Insurance and Access to Health Care

	Enrolled in public health insurance		Received medical care in the last 3 months: General practitioner		Emergency room		Specialist	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Reduced Form				
Yr Age 21	-0.008***		-0.004***		-0.001***		-0.003***	
	(0.001)		(0.001)		(0.000)		(0.001)	
Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)	0.009***	0.007***	-0.001	-0.003**	0.001	0.000	-0.001	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
				Panel B: IV				
𝟙(Any College)	-0.389***	-0.296***	0.050	0.106**	-0.025	-0.007	0.033	0.077**
	(0.046)	(0.049)	(0.050)	(0.047)	(0.019)	(0.019)	(0.032)	(0.033)
				Panel C: OLS				
𝟙(Any College)	-0.237***	-0.236***	0.007**	0.007**	-0.011***	-0.011***	0.031***	0.031***
	(0.006)	(0.006)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
County x Year x Gender FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	161,473	161,473	112,967	112,967	155,504	155,504	163,494	163,494
R-squared (panel A)	0.183	0.185	0.087	0.088	0.066	0.066	0.090	0.091
R-squared (panel C)	0.222	0.224	0.087	0.088	0.066	0.067	0.092	0.092
Mean DV	0.591	0.591	0.187	0.187	0.0557	0.0557	0.117	0.117
Kleibergen-Paap F-statistic (panel B)	374.4	331.4	237.4	197.2	349.2	319.7	378.6	341.5

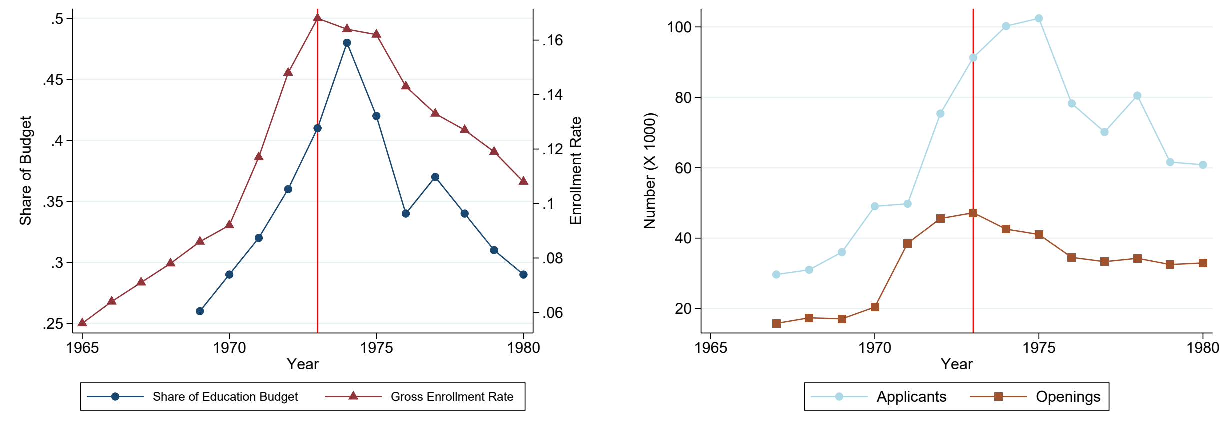
Notes: The unit of analysis is an individual respondent from the CASEN survey. Dependent variable in the header. A dummy for being enrolled in the public health insurance (FONASA) in columns 1-2; a dummy for having seen a general practitioner in the past three months in columns 3-4; similar dummies for visits to the emergency room or to a specialist in columns 5-6 and 7-8 respectively. Sample period: 1990-2017 (13 waves). Not all questions are asked in all years. Sample includes individuals from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and report 4+ years of secondary education. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. All regressions include year by county of residence by gender fixed effects. Even-numbered columns replace the baseline cohort trend with age fixed effects. Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Health Behaviors and Status

	Smoked in last month		Pap smear in last 3 years		Self-assessment: good health		Sick or in accident in last 3 months	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Reduced Form				
Yr Age 21	0.015***		0.019***		0.006***		-0.003***	
	(0.002)		(0.002)		(0.001)		(0.001)	
Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)	-0.014***	-0.009***	-0.012***	-0.004**	0.003**	0.002	0.000	-0.001
	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
				Panel B: IV				
𝟙(Any College)	0.587***	0.411***	0.498***	0.168**	-0.134**	-0.099	-0.012	0.033
	(0.114)	(0.148)	(0.095)	(0.084)	(0.065)	(0.080)	(0.032)	(0.036)
				Panel C: OLS				
𝟙(Any College)	-0.012	-0.014*	0.021***	0.020***	0.097***	0.097***	0.006**	0.006**
	(0.007)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)
County x Year x Gender FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes	No	Yes	No	Yes
Sample	All	All	Female	Female	All	All	All	All
Observations	29,613	29,613	58,549	58,549	86,706	86,706	162,416	162,416
R-squared (panel A)	0.058	0.059	0.103	0.111	0.106	0.107	0.102	0.102
R-squared (panel C)	0.057	0.059	0.103	0.111	0.112	0.113	0.102	0.102
Mean DV	0.397	0.397	0.726	0.726	0.526	0.526	0.179	0.179
Kleibergen-Paap F-statistic (panel B)	92.29	42.48	214.2	131.8	175.6	148.7	373.5	337.8

Notes: The unit of analysis is an individual respondent from the CASEN survey. Dependent variable in the header. A dummy for smoking in the past month in columns 1-2; a dummy for having a Pap smear in last three years in columns 3-4 (female only); a dummy for self-assessed health status good or very good (6-7 on 7-point scale or 4-5 on 5-point scale) in columns 6-7; a dummy for being sick or involved in an accident in the past 3 months in columns 7-8. Sample period: 1990-2017 (13 waves). Not all questions are asked in all years. Sample includes individuals from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and report 4+ years of secondary education. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. All regressions include year by county of residence by gender fixed effects. Even-numbered columns replace the baseline cohort trend with age fixed effects. Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Background Figures

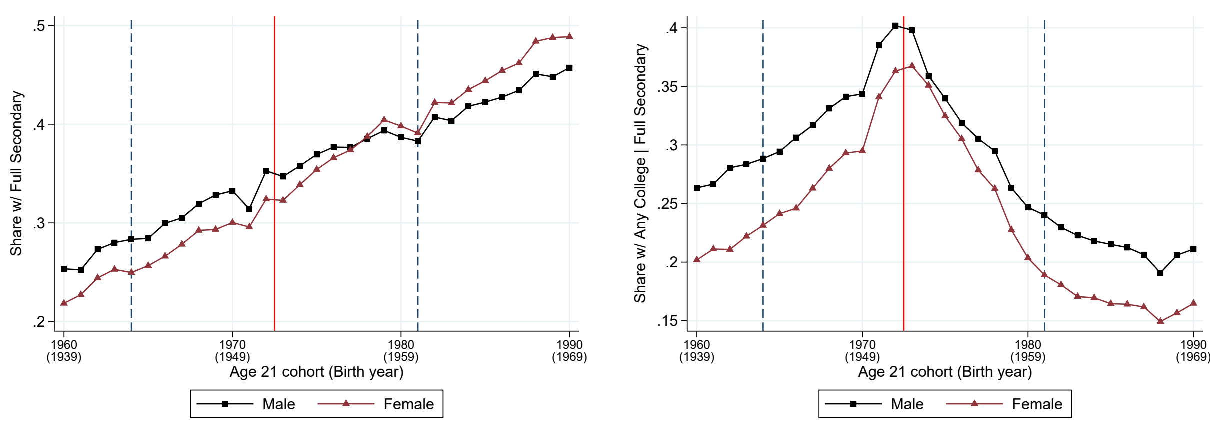


(a) Government subsidies and College Enrollment

(b) College Applicants and Openings

Note: Panel (a) shows the share of the government's education budget devoted to higher education (circle markers) and the gross enrollment rate in higher education (triangle markers). Panel (b) shows the yearly number of applicants for college (circle markers) and the number of openings on offer by the universities (square markers). Sources: Universidad de Chile (1972, 2011); PIIE (1984).

Figure 2: Educational Attainment by Cohort

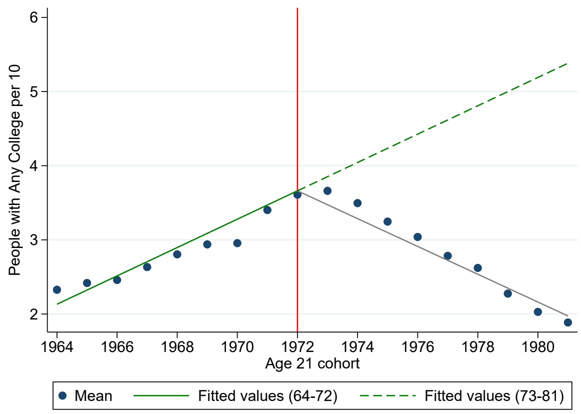


(a) Full Secondary

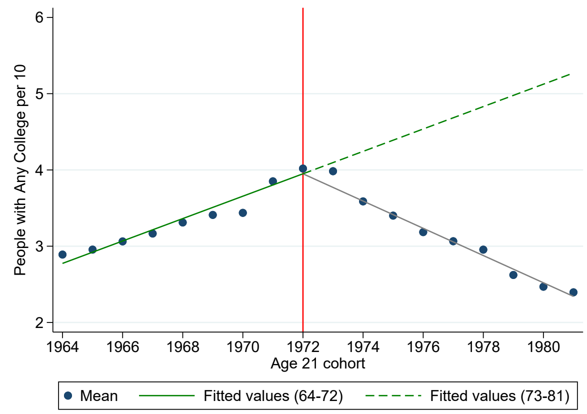
(b) Any College | Full Secondary

Notes: Panel (a) shows the shares of men and women per cohort (normalized to age 21) that report 4+ years of secondary in the 1992 census (our proxy for full secondary). Panel (b) shows the respective shares of people that report having any college education, among those with 4+ years of secondary. Dashed lines show the start (1964) and end date (1981) of the study cohorts included in the analysis.

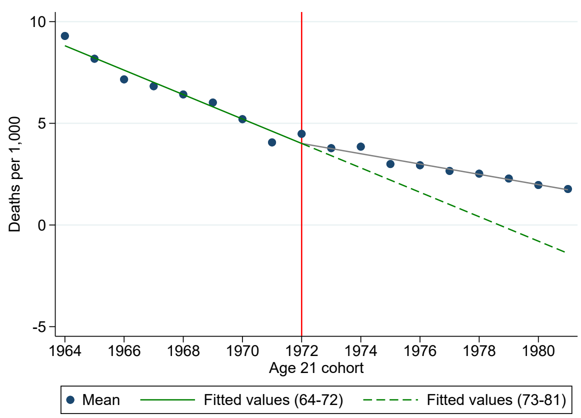
Figure 3: Visualization of Kink: College Enrollment and Mortality



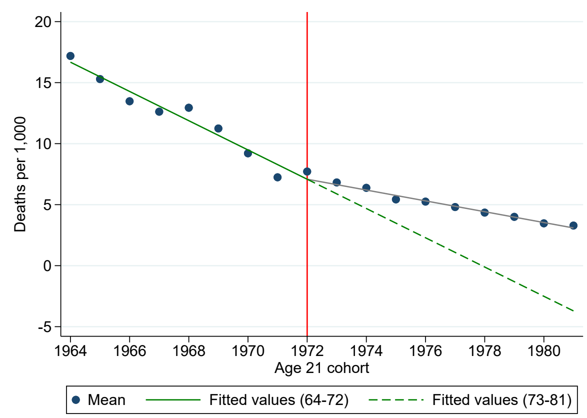
(a) Female: Any College



(b) Male: Any College



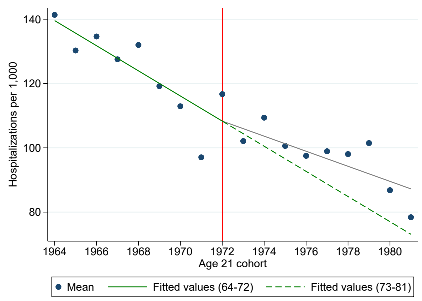
(c) Female: Mortality



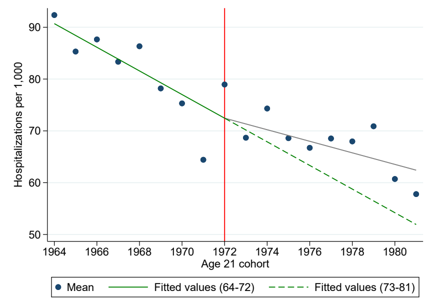
(d) Male: Mortality

Note: Panels show averages by cohort (across years) for the variable in the caption: Share with college per every 10 individuals in panels (a) and (b); Deaths per 1,000 individuals in panels (c) and (d). Both the mortality rate and the share with college are adjusted for previous mortality before averaging. Average is weighted by cell size (cohort-year). Solid green line corresponds to the line of best fit for cohorts reaching college age before 1973, which we extrapolate for later cohorts (dashed line). Grey line corresponds to line of best fit for cohorts reaching college age in 1973 or afterwards.

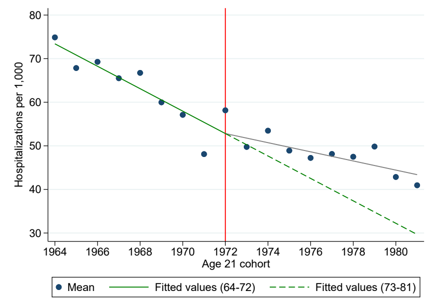
Figure 4: Visualization of Kink: Hospitalizations



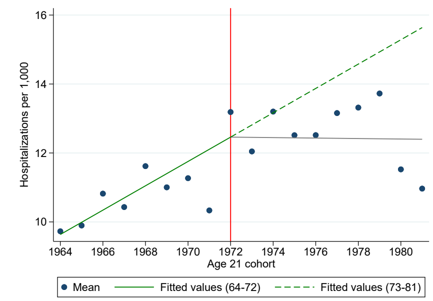
(a) Female: All admissions



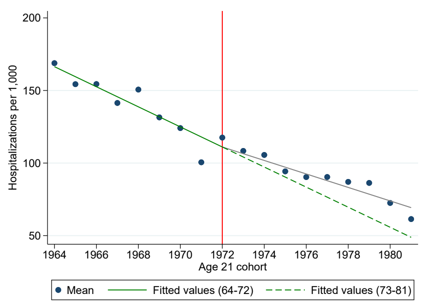
(b) Female: Admitted individuals



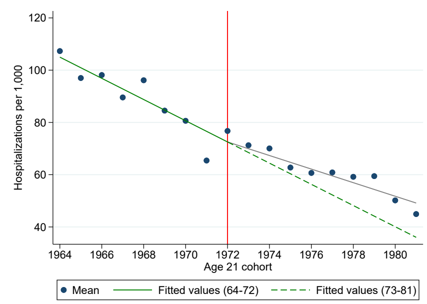
(c) Female: Public insurance



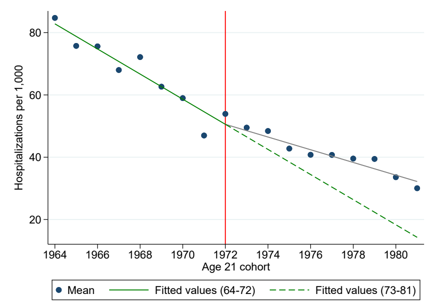
(d) Female: Private insurance



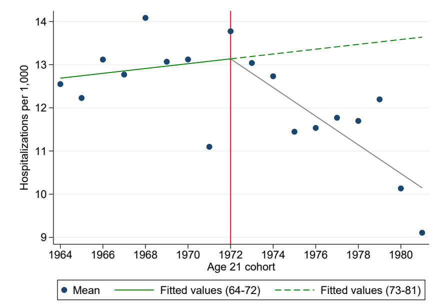
(e) Male: All admissions



(f) Male: Admitted individuals



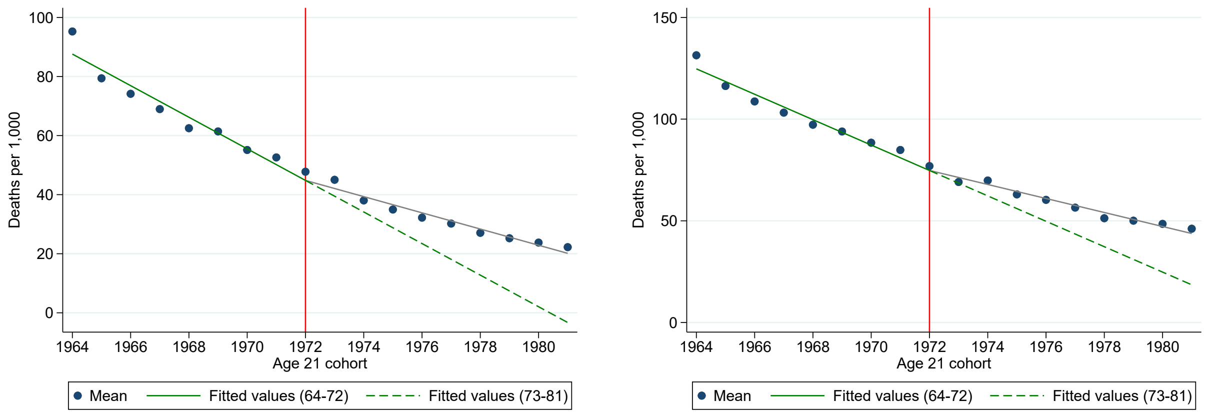
(g) Male: Public insurance



(h) Male: Private insurance

Note: Panels show averages by cohort (across years) for the variable in the caption. Panels (a) and (e) show the total number of hospital admissions (i.e. events) per 1,000. Panels (b) and (f) show the number of admitted patients per 1,000 (i.e. ignoring readmissions). Panels (c) and (g) show the number of admitted patients per 1,000 using public insurance (FONASA), while panels (d) and (h) shows patients using private insurance (ISAPRE). All hospitalization rates are adjusted for previous mortality before averaging. Average is weighted by cell size (cohort-year). Solid green line corresponds to the line of best fit for cohorts reaching college age before 1973, which we extrapolate for later cohorts (dashed line). Grey line corresponds to line of best fit for cohorts reaching college age in 1973 or afterwards.

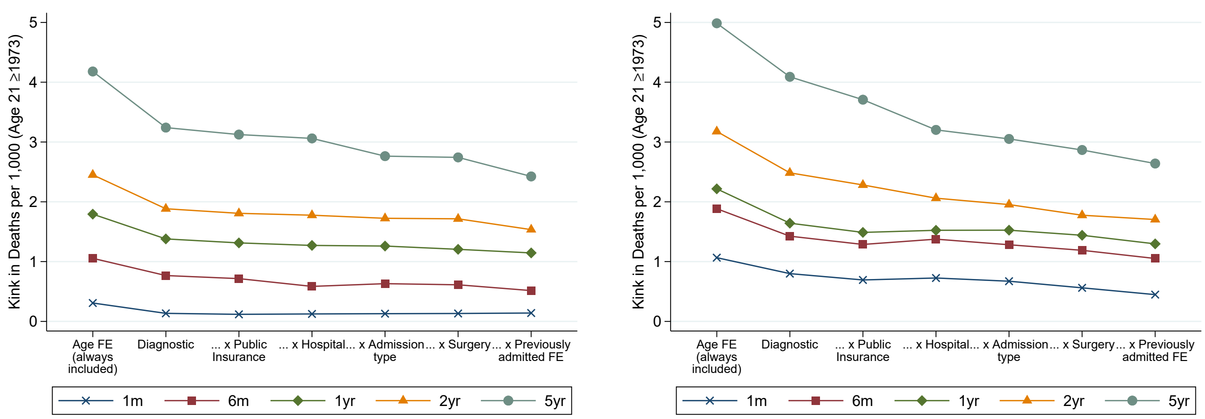
Figure 5: Visualization of Kink: One-year Mortality Rate of Hospitalized Patients



(a) Female **(b)** Male

Note: Panels show the average mortality rate within one year of hospital discharge by cohort (averaged across patients, one observation per patient corresponding to first admission). Sample includes individuals reaching age 21 between 1964 and 1981 (both inclusive). Sample period: 2004-2012. Solid green line corresponds to the line of best fit for cohorts reaching college age before 1973, which we extrapolate for later cohorts (dashed line). Grey line corresponds to line of best fit for cohorts reaching college age in 1973 or afterwards.

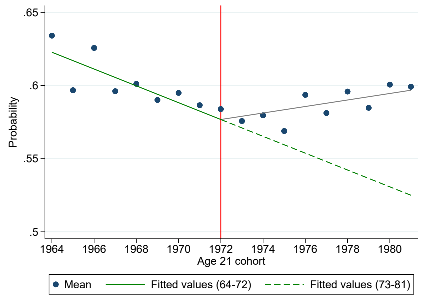
Figure 6: Mortality Rate of hospitalized patients at different time horizons



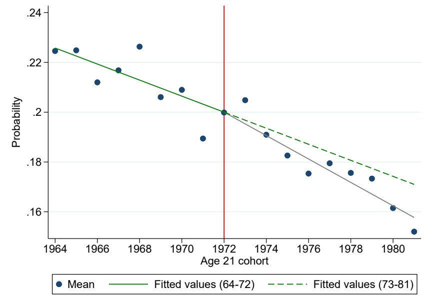
(a) Female **(b)** Male

Note: Figure plots the estimated size of the kink in mortality for hospitalized patients in cohorts reaching age 21 after 1973, i.e. coefficient for Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973), where "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972 and $\mathbb{1}$(Yr Age 21 $\geq$ 1973) is a dummy for cohorts that reached age 21 on or after 1973. Sample includes patients from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and is limited to one observation per patient (i.e. first admission). Sample period: 2004-2012. Different lines correspond to different time horizons for mortality relative to time of discharge. Different markers correspond to increasingly stringent sets of fixed effects, similarly to Table 5. Leftmost markers correspond to specification with age and county by year fixed effects, which are always included.

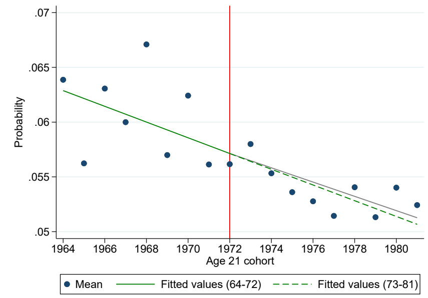
Figure 7: Visualization of Kink: Survey responses



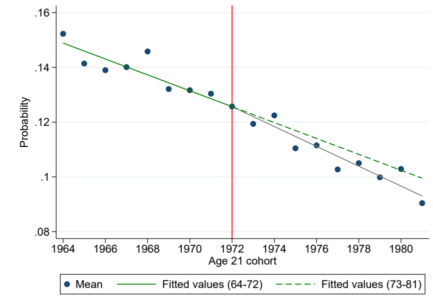
(a) Public insurance (FONASA)



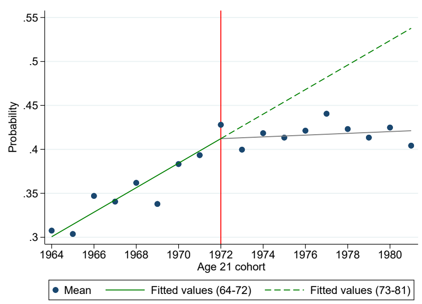
(b) GP in past 3 months



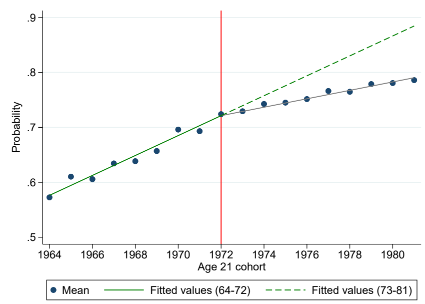
(c) ER in past 3 months



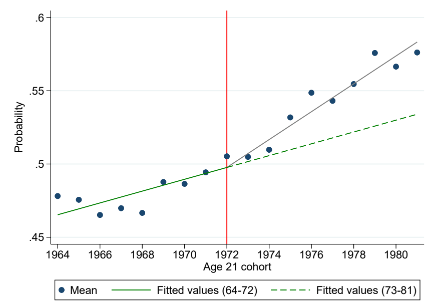
(d) Specialist in past 3 months



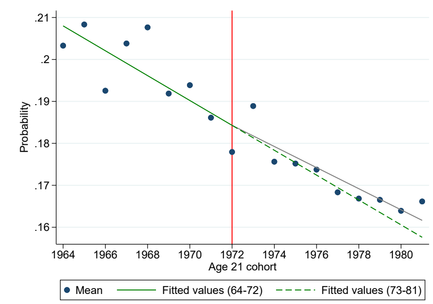
(e) Smoked in past month



(f) Pap smear in past 3 years



(g) Good self-assessed health



(h) Sick or accident in past 3 months

Note: Panels show averages by cohort (across years) for the variable in the caption, based on individual responses to the CASEN survey. In panel (a), the enrollment rate in public insurance (FONASA). In panels (b)-(d), the probability of visiting a general practitioner, the emergency room or a specialist in the past three months. In panel (e), the probability of having smoked in the past month. In panel (f), the probability of having had a Pap smear in the past three years (female only). In panel (g), the probability of reporting good or very good health (4-5 on 5-point scale or 6-7 in 7-point scale). In panel (h), the probability of being sick or in an accident in the past three months. Solid green line corresponds to the line of best fit for cohorts reaching college age before 1973, which we extrapolate for later cohorts (dashed line). Grey line corresponds to line of best fit for cohorts reaching college age in 1973 or afterwards.

Appendix (for online publication)

Table of Contents

Appendix A Additional Results: Mortality

Table A1: First Stage: Using 2002 Census

	Female			Male		
	Adj share w/ college per 10 people	Share w/ college per 10 people	Average years of college	Adj share w/ college per 10 people	Share w/ college per 10 people	Average years of college
	(1)	(2)	(3)	(4)	(5)	(6)
Yr Age 21	0.140***	0.142***	0.048***	0.010***	0.100***	0.030***
	(0.004)	(0.011)	(0.003)	(0.000)	(0.007)	(0.003)
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	-0.277***	-0.276***	-0.093***	-0.023***	-0.225***	-0.071***
	(0.007)	(0.018)	(0.006)	(0.000)	(0.012)	(0.005)
Year x Region FE	Yes	No	No	Yes	No	No
Region FE	No	Yes	Yes	No	Yes	Yes
Adjusted for deaths	Yes	No	No	Yes	No	No
Observations	4,050	270	270	4,050	270	270
R-squared	0.832	0.835	0.835	0.884	0.883	0.859
Mean DV	3.004	2.978	1.055	0.351	3.516	1.298

Notes: The unit of analysis is cohort-region-year of death in columns 1 and 4, and cohort-region in columns 2, 3, 5, and 6. Dependent variable in the header. “Any College” is the number of people that enrolled in college out of every ten. Original sample includes all respondents of the 2002 census from cohorts born between 1943 and 1960 (both inclusive) that report full secondary education. “Yr Age 21” is a continuous variable indicating the year at which the cohort reached 21 years of age, normalized to zero in 1972. “Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)” is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. All regressions include region fixed effects, columns 1 and 4 include Year of Death x Region fixed effects. Standard errors clustered by region-year in parentheses in columns 1 and 4, robust standard errors otherwise. *** p<0.01, ** p<0.05, * p<0.1

Table A2: Mortality Rate: Years of college

	Dependent variable: Deaths per 1,000 people			
	Female		Male	
	Trend	Age FE	Trend	Age FE
	(1)	(2)	(3)	(4)
	Panel A: IV			
Years of college	-2.282***	-0.687***	-6.164***	-2.415***
	(0.150)	(0.142)	(0.368)	(0.319)
	Panel B: OLS			
Years of college	-1.893***	-0.508***	-5.176***	-2.176***
	(0.119)	(0.120)	(0.256)	(0.268)
Year x Region FE	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes
Observations	6,480	6,480	6,480	6,480
R-squared (panel B)	0.673	0.755	0.745	0.850
Mean DV	3.850	3.850	7.084	7.084
Kleibergen-Paap F-statistic (panel A)	13215	10390	12299	9333

Notes: Sample includes all respondents in the 1992 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) reported full secondary education. Sample period: 1994-2017. In panel A, the interaction between a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972, and a dummy for cohorts that reached age 21 on or after 1973 is used as excluded instrument for the average years of college. Mortality rate is adjusted for previous mortality. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Mortality Rate: Robustness to unadjusted share of college

	Dependent variable: Deaths per 1,000 people			
	Female		Male	
	Trend	Age FE	Trend	Age FE
	(1)	(2)	(3)	(4)
	Panel A: IV			
Share with college per 10 people	-0.893***	-0.269***	-2.416***	-0.948***
	(0.059)	(0.056)	(0.138)	(0.125)
	Panel B: OLS			
Share with college per 10 people	-0.733***	-0.194***	-2.111***	-0.873***
	(0.048)	(0.046)	(0.107)	(0.106)
Year x Region FE	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes
Observations	6,480	6,480	6,480	6,480
R-squared (panel B)	0.673	0.755	0.746	0.850
Mean DV	3.850	3.850	7.084	7.084
Kleibergen-Paap F-statistic (panel A)	5608	4990	10918	9842

Notes: Unit of analysis is cohort-region-year. Observations weighted by cell size. Sample includes all respondents in the 1992 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) reported full secondary education. Sample period: 1994-2017. In panel A, the interaction between a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972, and a dummy for cohorts that reached age 21 on or after 1973 is used as excluded instrument for the share with college. Mortality rate is adjusted for previous mortality. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Mortality Rate: Robustness to unadjusted mortality

	Dependent variable: Deaths per 1,000 people			
	Female		Male	
	Trend	Age FE	Trend	Age FE
	(1)	(2)	(3)	(4)
	Panel A: Reduced Form			
Yr Age 21	-0.529***		-0.961***	
	(0.024)		(0.038)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	0.289***	0.103***	0.560***	0.282***
	(0.019)	(0.020)	(0.029)	(0.033)
	Panel B: IV			
Share with college per 10 people	-0.715***	-0.253***	-1.687***	-0.825***
	(0.045)	(0.049)	(0.087)	(0.094)
	Panel C: OLS			
Share with college per 10 people	-0.554***	-0.183***	-1.203***	-0.715***
	(0.038)	(0.042)	(0.075)	(0.080)
Year x Region FE	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes
Observations	6,480	6,480	6,480	6,480
R-squared (panel A)	0.679	0.737	0.770	0.831
R-squared (panel C)	0.675	0.737	0.762	0.830
Mean DV	3.644	3.644	6.402	6.402
Kleibergen-Paap F-statistic (panel B)	5015	4574	9875	8783

Unit of analysis is cohort-region-year. Observations weighted by cell size. Sample includes all respondents in the 1992 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) reported full secondary education. Sample period: 1994-2017. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. The share with college is adjusted for previous mortality. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Mortality Rate: Robustness to unadjusted mortality rate and share of college

	Dependent variable: Deaths per 1,000 people			
	Female		Male	
	Trend	Age FE	Trend	Age FE
	(1)	(2)	(3)	(4)
	Panel A: Reduced Form			
Yr Age 21	-0.529***		-0.961***	
	(0.024)		(0.038)	
Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)	0.289***	0.103***	0.560***	0.282***
	(0.019)	(0.020)	(0.029)	(0.033)
	Panel B: IV			
Share with college per 10 people	-0.725***	-0.259***	-1.692***	-0.848***
	(0.046)	(0.050)	(0.087)	(0.097)
	Panel C: OLS			
Share with college per 10 people	-0.595***	-0.187***	-1.481***	-0.761***
	(0.039)	(0.042)	(0.072)	(0.083)
Year x Region FE	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes
Observations	6,480	6,480	6,480	6,480
R-squared (panel A)	0.679	0.737	0.770	0.831
R-squared (panel C)	0.676	0.737	0.767	0.831
Mean DV	3.644	3.644	6.402	6.402
Kleibergen-Paap F-statistic (panel B)	5608	4990	10918	9842

Unit of analysis is cohort-region-year. Observations weighted by cell size. Sample includes all respondents in the 1992 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) reported full secondary education. Sample period: 1994-2017. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Mortality Rate: National level

	Dependent variable: Deaths per 1,000 people			
	Female		Male	
	Trend	Age FE	Trend	Age FE
	(1)	(2)	(3)	(4)
	Panel A: Reduced Form			
Yr Age 21	-0.605***		-1.232***	
	(0.062)		(0.128)	
Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)	0.349***	0.102***	0.784***	0.302***
	(0.039)	(0.018)	(0.080)	(0.043)
	Panel B: IV			
Share with college per 10 people	-0.858***	-0.248***	-2.358***	-0.883***
	(0.094)	(0.044)	(0.239)	(0.124)
	Panel C: OLS			
Share with college per 10 people	-0.742***	-0.214***	-1.795***	-0.852***
	(0.080)	(0.042)	(0.194)	(0.119)
Year FE	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes
Observations	432	430	432	430
R-squared (panel A)	0.866	0.971	0.858	0.979
R-squared (panel C)	0.862	0.971	0.848	0.979
Mean DV	3.849	3.847	7.077	7.069
Kleibergen-Paap F-statistic (panel B)	110038	26971	340546	61740

Notes: Unit of analysis is cohort-year. Observations weighted by cell size. Sample includes all respondents in the 1992 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) reported full secondary education. Sample period: 1994-2017. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. Mortality rate and share with college adjusted for previous mortality. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1. P-value from Wild cluster bootstrap by cohort reported in brackets.

Table A7: Mortality Rate: Using 2002 Census

	Dependent variable: Deaths per 1,000 people			
	Female		Male	
	Trend	Age FE	Trend	Age FE
	(1)	(2)	(3)	(4)
	Panel A: Reduced Form			
Yr Age 21	-0.729***		-1.285***	
	(0.034)		(0.053)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	0.486***	0.207***	0.861***	0.415***
	(0.030)	(0.032)	(0.045)	(0.053)
	Panel B: IV			
Share with college per 10 people	-1.752***	-0.744***	-3.907***	-1.796***
	(0.115)	(0.114)	(0.220)	(0.232)
	Panel C: OLS			
Share with college per 10 people	-1.129***	-0.336***	-2.319***	-0.982***
	(0.081)	(0.087)	(0.150)	(0.168)
Year x Region FE	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes
Observations	4,050	4,050	4,050	4,050
R-squared (panel A)	0.673	0.718	0.780	0.826
R-squared (panel C)	0.658	0.714	0.759	0.821
Mean DV	4.395	4.395	7.436	7.436
Kleibergen-Paap F-statistic (panel B)	1740	1524	1915	1973

Notes: Unit of analysis is cohort-region-year. Observations weighted by cell size. Sample includes all respondents in the 2002 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) reported full secondary education. Sample period: 2003-2017. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. Mortality rate and share with college adjusted for previous mortality. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Marriage, widow status, and education

	Female		Male	
	Ever married	Widow	Ever married	Widow
	(1)	(2)	(3)	(4)
	Panel A: Reduced Form			
Yr Age 21	0.0001	-0.004***	-0.0004*	-0.001***
	(0.0003)	(0.0002)	(0.0002)	(0.0001)
Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)	-0.008***	0.003***	-0.011***	0.001***
	(0.0004)	(0.0002)	(0.0004)	(0.0001)
	Panel B: IV			
$\mathbb{1}$(Any College)	0.200***	-0.079***	0.330***	-0.017***
	(0.012)	(0.005)	(0.014)	(0.003)
County FE	Yes	Yes	Yes	Yes
Married FE	No	Yes	No	Yes
Observations	514,886	514,886	509,684	509,684
R-squared (panel A)	0.010	0.014	0.017	0.003
Mean DV	0.816	0.0165	0.895	0.0036
Kleibergen-Paap F-statistic (panel B)	2297.4	2332.0	1899.2	1934.5

Notes: Sample includes all respondents in the 1992 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) that reported full secondary education. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. All regressions include County FE. Standard errors clustered by county in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A9: Cause of Death

	Dependent variable: Deaths per 1,000 people				
	Tumors	Circulatory system	External causes	Digestive system	Respiratory system
	(1)	(2)	(3)	(4)	(5)
	Panel A: Female (Linear trend)				
Share with college per 10 people	-0.342***	-0.235***	-0.005	-0.065***	-0.068***
	(0.030)	(0.023)	(0.009)	(0.011)	(0.011)
	Panel B: Female (Age FE)				
Share with college per 10 people	-0.108***	-0.083***	0.001	-0.033***	-0.006
	(0.033)	(0.021)	(0.010)	(0.011)	(0.009)
	Panel C: Male (Linear trend)				
Share with college per 10 people	-0.839***	-0.711***	-0.078***	-0.225***	-0.182***
	(0.058)	(0.051)	(0.025)	(0.024)	(0.020)
	Panel D: Male (Age FE)				
Share with college per 10 people	-0.213***	-0.293***	-0.064***	-0.173***	-0.059***
	(0.051)	(0.052)	(0.024)	(0.025)	(0.015)
Year x Region FE	Yes	Yes	Yes	Yes	Yes
Observations	6,480	6,480	6,480	6,480	6,480
Mean DV (Female)	1.974	0.713	0.223	0.263	0.164
Mean DV (Male)	2.229	1.915	0.883	0.779	0.332
Kleibergen-Paap F-statistic (panel A)	5015	5015	5015	5015	5015
Kleibergen-Paap F-statistic (panel B)	4574	4574	4574	4574	4574
Kleibergen-Paap F-statistic (panel C)	9875	9875	9875	9875	9875
Kleibergen-Paap F-statistic (panel D)	8783	8783	8783	8783	8783

Notes: Unit of analysis is cohort-region-year. Observations weighted by cell size. Sample includes all respondents in the 1992 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) reported full secondary education. Sample period: 1994-2017. Mortality rates and share with college adjusted for previous mortality. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. This interaction is used as excluded instrument for the share with college. In panels A and C the baseline trend is included as additional control, while in panels B and D it is replaced by age fixed effects. All regressions include Year x Region FE. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A10: Cause of Death: Disaggregated Tumors

	Dependent variable: Deaths per 1,000 people				
	(1)	(2)	(3)	(4)	(5)
	Digestive	**Breast**	**Genital**	**Respiratory**	**Lymphatic**
	Panel A: Female (Linear trend)				
Share with college per 10 people	-0.134***	-0.032***	-0.039***	-0.040***	-0.050***
	(0.017)	(0.012)	(0.009)	(0.012)	(0.008)
	Panel B: Female (Age FE)				
Share with college per 10 people	-0.042**	-0.030**	-0.023**	0.015	-0.027***
	(0.019)	(0.012)	(0.010)	(0.012)	(0.008)
	Digestive	**Respiratory**	**Lymphatic**	**Urinary**	**Genital**
	Panel C: Male (Linear trend)				
Share with college per 10 people	-0.326***	-0.197***	-0.084***	-0.061***	-0.088***
	(0.029)	(0.020)	(0.014)	(0.012)	(0.014)
	Panel D: Male (Age FE)				
Share with college per 10 people	-0.092***	-0.060***	-0.023**	-0.008	0.003
	(0.031)	(0.019)	(0.011)	(0.011)	(0.009)
Year x Region FE	Yes	Yes	Yes	Yes	Yes
Observations	6,480	6,480	6,480	6,480	6,480
Mean DV (Female)	0.633	0.367	0.295	0.228	0.147
Mean DV (Male)	0.906	0.437	0.208	0.169	0.137
Kleibergen-Paap F-statistic (panel A)	5015	5015	5015	5015	5015
Kleibergen-Paap F-statistic (panel B)	4574	4574	4574	4574	4574
Kleibergen-Paap F-statistic (panel C)	9875	9875	9875	9875	9875
Kleibergen-Paap F-statistic (panel D)	8782	8782	8782	8782	8782

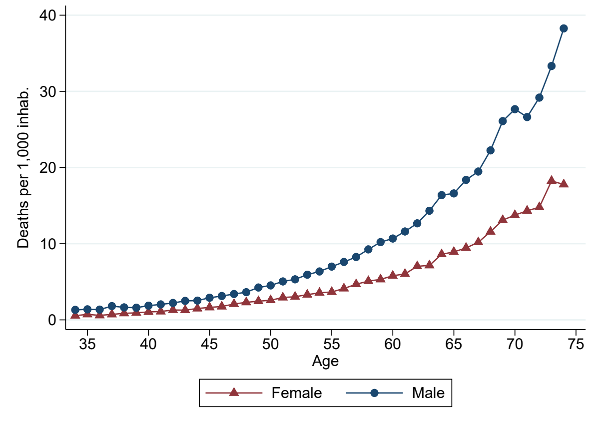
Notes: Unit of analysis is cohort-region-year. Observations weighted by cell size. Sample includes all respondents in the 1992 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) reported full secondary education. Sample period: 1994-2017. Mortality rates and share with college adjusted for previous mortality. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. This interaction is used as excluded instrument for the share with college. In panels A and C the baseline trend is included as additional control, while in panels B and D it is replaced by age fixed effects. All regressions include Year x Region FE. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A11: Cause of Death: Disaggregated External Causes

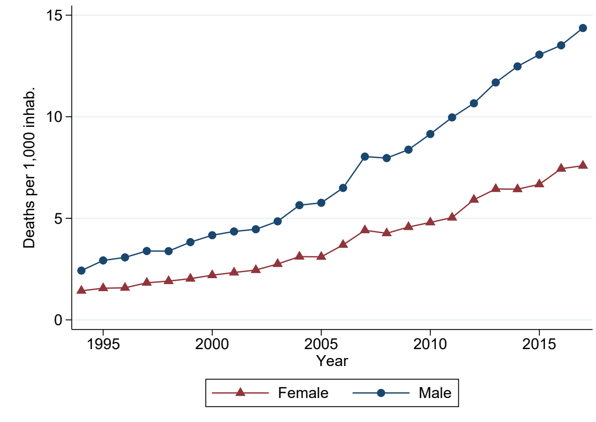
	Transit accident	Other accident	Homicide	Suicide	Medical complication	Other
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Female (Linear trend)			
Share with college per 10 people	-0.001	-0.004	-0.003	0.003	-0.001	0.000
	(0.006)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
			Panel B: Female (Age FE)			
Share with college per 10 people	0.003	0.003	-0.002	-0.001	-0.001	-0.001
	(0.006)	(0.005)	(0.002)	(0.004)	(0.002)	(0.004)
			Panel C: Male (Linear trend)			
Share with college per 10 people	-0.028**	-0.047***	0.001	0.011	-0.006***	-0.009
	(0.013)	(0.013)	(0.005)	(0.011)	(0.002)	(0.007)
			Panel D: Male (Age FE)			
Share with college per 10 people	-0.020	-0.030**	-0.003	0.014	-0.005**	-0.021***
	(0.014)	(0.012)	(0.005)	(0.012)	(0.002)	(0.006)
Year x Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,480	6,480	6,480	6,480	6,480	6,480
Mean DV (Female)	0.0733	0.0540	0.0109	0.0567	0.0050	0.0224
Mean DV (Male)	0.284	0.242	0.0454	0.219	0.0066	0.0856
Kleibergen-Paap F-statistic (panel A)	5015	5015	5015	5015	5015	5015
Kleibergen-Paap F-statistic (panel B)	4574	4574	4574	4574	4574	4574
Kleibergen-Paap F-statistic (panel C)	9875	9875	9875	9875	9875	9875
Kleibergen-Paap F-statistic (panel D)	8783	8783	8783	8783	8783	8783

Notes: Unit of analysis is cohort-region-year. Observations weighted by cell size. Sample includes all respondents in the 1992 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) reported full secondary education. Sample period: 1994-2017. Mortality rates and share with college adjusted for previous mortality. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. This interaction is used as excluded instrument for the share with college. In panels A and C the baseline trend is included as additional control, while in panels B and D it is replaced by age fixed effects. All regressions include Year x Region FE. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

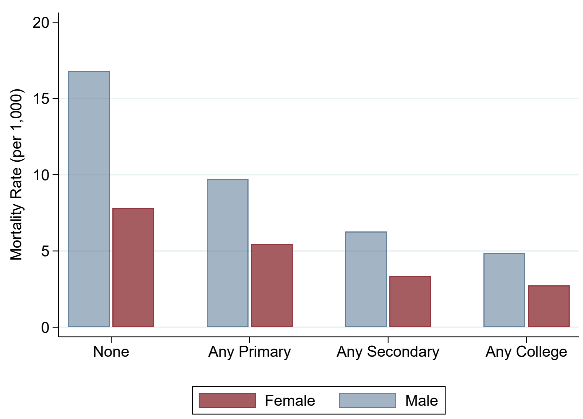
Figure A1: Mortality Rate by Age, Year and Education



(a) Age



(b) Year



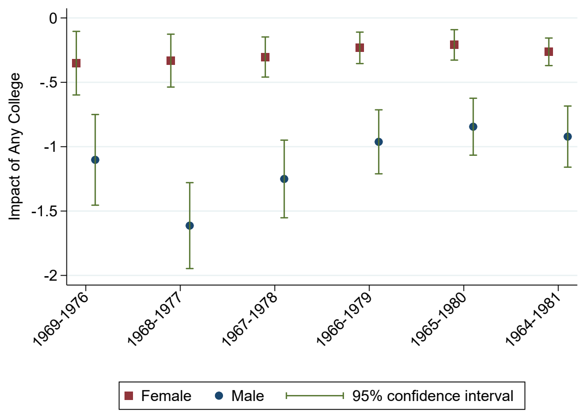
(c) Educational level

Notes: Panels (a) and (b) show female and male mortality rates by age and by year, respectively. Sample period: 1994-2017. Mortality rates calculated based on population counts in the 1992 census, only including individuals from cohorts reaching age 21 between 1964 and 1981 (both inclusive) and reporting 4+ years of secondary education. These mortality rates have been adjusted for previous mortality per cell (cohort). Panel (c) shows mortality rates by highest level of education for the same period. This sample does not impose any restriction on educational attainment.

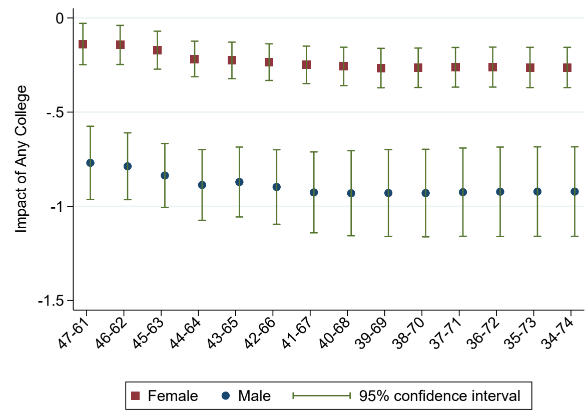
Figure A2: Age distribution of Study Cohorts

Notes: This figure shows the number of cohorts observed at each age. We disaggregate between cohorts reaching age 21 before 1973 (i.e. age 21 between 1964 and 1972) and those reaching the same age on or after 1973 (i.e. age 21 between 1973 and 1981). Sample period: 1994-2017.

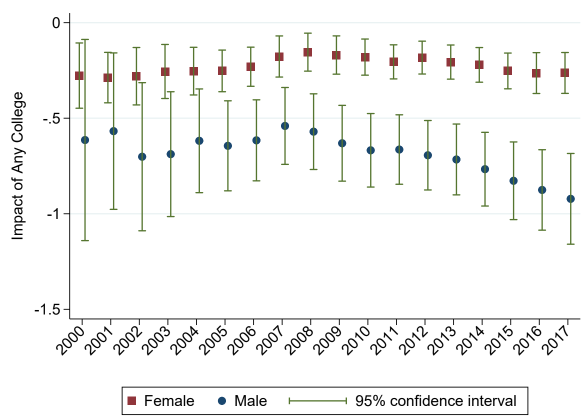
Figure A3: Robustness: Mortality w/ Different Bandwidths



(a) Cohorts



(b) Ages



(c) End year

Notes: Each panel replicates the IV estimation of the impact of college on mortality for varying bandwidths, along different dimensions. The baseline sample includes cohorts reaching age 21 between 1964 and 1981 and the sample period is 1994-2017, without any age restrictions. In panel (a) we consider different bandwidths for the set of cohorts included in the sample. In panel (b), we vary the set of ages at which we allow cohorts to enter the sample. In panel (c), we modify the final year of the sample. In all cases, sample is restricted to people with full secondary education. Dependent variable is the number of deaths per 1,000 individuals, while the endogenous variable is the share with college per 10 individuals. The excluded instrument is the interaction of "Yr Age 21", a continuous variable indicating the year at which the cohort reached 21 years of age (normalized to zero in 1972) with a dummy for cohorts that reached age 21 on or after 1973. All regressions include region by year and age fixed effects (i.e. equation 6). Standard errors clustered by region-year. Bars correspond to 95% confidence interval.

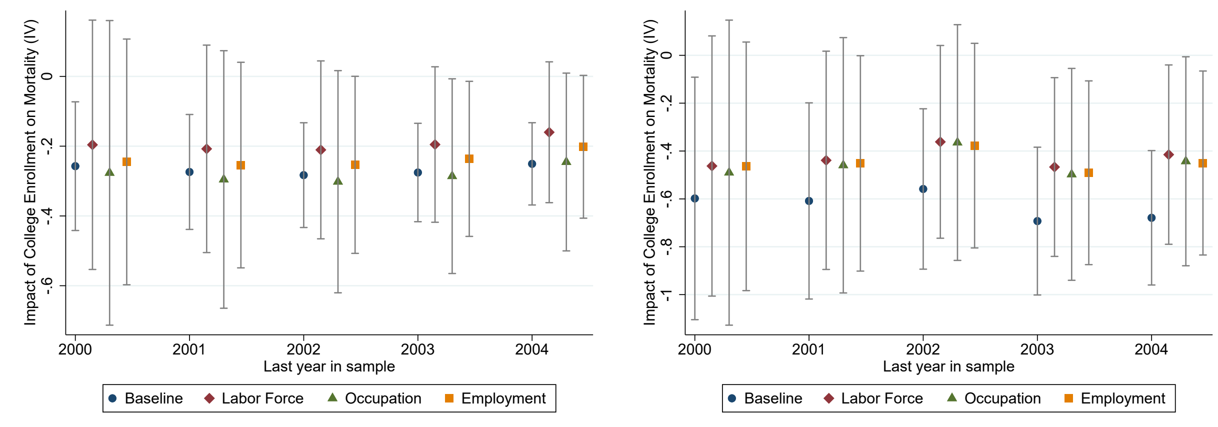
Appendix B Additional Results: Labor Market Outcomes

Table B1: Summary Statistics: Labor Market Outcomes

	Mean	SD	Min	Max
	(1)	(2)	(3)	(4)
Panel A: Female (N=4,492,512)				
Labor Force Participation:				
In labor force	0.572	0.022	0.538	0.612
Domestic work	0.396	0.021	0.361	0.430
Other activities	0.031	0.007	0.026	0.056
Occupation category:				
White collar, high skill	0.265	0.037	0.200	0.326
White collar, low skill	0.220	0.014	0.194	0.240
Blue collar	0.054	0.005	0.047	0.065
Military	0.004	0.001	0.002	0.006
Unemployed/inactive (also below)	0.455	0.025	0.414	0.497
Employment type:				
Business owner	0.045	0.006	0.035	0.055
Salaried employee	0.446	0.019	0.409	0.484
Self-employed	0.054	0.006	0.043	0.065
Panel B: Male (N=4,490,721)				
Labor Force Participation:				
In labor force	0.948	0.014	0.901	0.959
Domestic work	0.006	0.001	0.005	0.007
Other activities	0.046	0.014	0.036	0.093
Occupation category:				
White collar, high skill	0.365	0.033	0.310	0.415
White collar, low skill	0.247	0.005	0.236	0.255
Blue collar	0.248	0.023	0.216	0.288
Military	0.046	0.015	0.018	0.067
Unemployed/inactive (also below)	0.090	0.015	0.075	0.137
Employment type:				
Business owner	0.103	0.015	0.081	0.128
Salaried employee	0.667	0.037	0.580	0.704
Self-employed	0.140	0.014	0.116	0.165

Notes: Table shows summary statistics by gender and labor market outcome. Source: 1992 census.

Figure B1: Robustness to the last year of the sample



(a) Female **(b)** Male

Notes: Figure replicates the analysis in Table 3 for different sample end-years. In all cases, sample start year is 1994. Notes: The unit of analysis is cohort-year in the baseline estimation and cohort-region-category in all others. Observations are weighted by cell size. Categories determined by the respective labor market outcome. Labor force: In labor force, domestic duties, other activities. Occupation: white-collar high-skill, white-collar low-skill, blue-collar, military, unemployed/inactive. Employment: business owner, employee, self-employed, unemployed/inactive. Population at risk includes all respondents in the 1992 census who (I) reached age 21 between 1964 and 1981 (both years inclusive) and (II) reported full secondary education. The interaction of “Yr Age 21”, a continuous variable indicating the year at which the cohort reached age 21 (normalized to zero in 1972) with a dummy for cohorts that reached age 21 on or after 1973 is used as excluded instrument for the share with college. Mortality rate and share with college adjusted for previous mortality. All regressions include year and age fixed effects. 95% confidence intervals based on robust standard errors.

Appendix C Additional Results: Hospitalized patients

Table C1: Hospitalizations: Readmissions and Missing Patient IDs

	Readmissions		Admissions w/o ID	
	(1)	(2)	(3)	(4)
		Panel A: Female		
Yr Age 21	-1.319***		-0.269***	
	(0.058)		(0.043)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	0.149**	-0.410***	0.174***	0.383***
	(0.060)	(0.068)	(0.044)	(0.064)
		Panel B: Male		
Yr Age 21	-2.309***		-0.562***	
	(0.088)		(0.082)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	0.506***	-0.522***	0.280***	0.798***
	(0.063)	(0.109)	(0.050)	(0.115)
Year x Region FE	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes
Observations	4,590	4,590	4,590	4,590
R-squared (panel A)	0.819	0.842	0.967	0.969
R-squared (panel B)	0.861	0.894	0.917	0.933
Mean DV (panel A)	29.73	29.73	9.755	9.755
Mean DV (panel B)	34.35	34.35	8.932	8.932

Notes: The unit of analysis is cohort-region-year. Observations weighted by cell size. Dependent variable in the header. Number of hospital readmissions (per patient and year) in columns 1-2. Number of admissions without a patient ID in columns 3-4. Hospitalization rates adjusted for previous mortality within cell. Population at risk includes all respondents in the 2002 census that reached age 21 between 1964 and 1981 (both years inclusive). Sample period: 2002-2018. "Yr Age 21" is a continuous variable indicating the year when the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C2: Mortality Rate: Unrestricted Sample (2002-2017)

	Dependent variable: Deaths per 1,000 people			
	Female		Male	
	Trend	Age FE	Trend	Age FE
	(1)	(2)	(3)	(4)
	Panel A: Reduced Form			
Yr Age 21	-0.841***		-1.413***	
	(0.035)		(0.050)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)	0.509***	0.124***	0.871***	0.286***
	(0.024)	(0.017)	(0.034)	(0.029)
	Panel B: IV			
Share with college per 10 people	-3.893***	-0.950***	-6.562***	-2.083***
	(0.191)	(0.130)	(0.278)	(0.208)
	Panel C: OLS			
Share with college per 10 people	-2.292***	-0.227**	-3.395***	-0.571***
	(0.105)	(0.092)	(0.165)	(0.162)
Year x Region FE	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes
Observations	4,320	4,320	4,320	4,320
R-squared (panel A)	0.830	0.905	0.853	0.922
R-squared (panel C)	0.815	0.904	0.832	0.919
Mean DV	5.385	5.385	9.378	9.378
Kleibergen-Paap F-statistic (panel B)	4699	3020	4946	3162

Notes: Sample includes all respondents in the 2002 census that reached age 21 between 1964 and 1981 (both years inclusive), but imposes no restriction on educational attainment. Sample period: 2002-2017. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. Mortality rate and share with college adjusted for previous mortality. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table C3: Hospitalizations: Main diagnostic

	Tumors	Circulatory	External	Digestive	Respiratory	Musculoskeletal	Genitourinary	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Female (Trend)				
Yr Age 21	-0.313***	-1.163***	-0.278***	-0.363***	-0.725***	-0.019	-0.180***	-0.841***
	(0.038)	(0.036)	(0.019)	(0.018)	(0.034)	(0.017)	(0.035)	(0.038)
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	0.207***	0.498***	0.067***	-0.096***	0.331***	-0.266***	0.159***	0.531***
	(0.071)	(0.026)	(0.022)	(0.027)	(0.024)	(0.023)	(0.050)	(0.080)
				Panel B: Female (Age fixed effects)				
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	-0.096**	0.176***	-0.045**	0.003	0.076*	-0.217***	-0.108***	0.116**
	(0.042)	(0.042)	(0.019)	(0.035)	(0.041)	(0.026)	(0.035)	(0.052)
				Panel C: Male (Trend)				
Yr Age 21	-1.138***	-1.790***	-0.043**	-0.812***	-0.961***	0.014	-1.000***	-1.327***
	(0.034)	(0.044)	(0.017)	(0.022)	(0.041)	(0.017)	(0.027)	(0.047)
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	0.330***	0.578***	0.034	0.148***	0.570***	-0.149***	0.362***	0.410***
	(0.043)	(0.036)	(0.025)	(0.030)	(0.030)	(0.024)	(0.037)	(0.040)
				Panel D: Male (Age fixed effects)				
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	-0.075*	0.093	-0.012	0.096***	0.177***	-0.134***	0.053*	0.113
	(0.043)	(0.061)	(0.029)	(0.035)	(0.038)	(0.024)	(0.032)	(0.073)
Year x Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,590	4,590	4,590	4,590	4,590	4,590	4,590	4,590
R-squared (panel A)	0.451	0.852	0.670	0.688	0.778	0.761	0.505	0.715
R-squared (panel B)	0.601	0.874	0.689	0.700	0.808	0.769	0.643	0.778
R-squared (panel C)	0.829	0.888	0.700	0.798	0.779	0.665	0.852	0.825
R-squared (panel D)	0.860	0.908	0.705	0.801	0.842	0.674	0.885	0.835
Mean DV (Female)	16.33	14.48	9.262	19.91	8.287	8.155	14.67	27.90
Mean DV (Male)	12.99	21.80	13.67	20.14	8.880	6.977	12.23	26.82

Notes: The unit of analysis is cohort-region-year. Observations weighted by cell size. The dependent variable in each column is the number of hospitalizations per 1,000 (including readmissions) with diagnostic corresponding to the chapter from the ICD-10 classification in the header. Hospitalization rates adjusted for previous mortality within cell. Population at risk includes all respondents in the 2002 census that reached age 21 between 1964 and 1981 (both years inclusive). Sample period: 2002-2018. "Yr Age 21" is a continuous variable indicating the year when the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. Panels B and D replace the baseline cohort trend with age fixed effects. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C4: Hospitalizations: Other insurance categories

	Other insurance		No insurance		Insurance N/A	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Female			
Yr Age 21	-0.035***		0.012**		-0.024**	
	(0.010)		(0.005)		(0.010)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)	-0.015*	-0.079***	-0.007	0.013*	-0.004	0.046***
	(0.008)	(0.014)	(0.006)	(0.007)	(0.014)	(0.016)
			Panel B: Male			
Yr Age 21	-0.059***		-0.011		0.010	
	(0.012)		(0.007)		(0.021)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)	-0.059***	-0.187***	-0.014*	0.010	-0.118***	-0.030
	(0.012)	(0.021)	(0.008)	(0.009)	(0.024)	(0.042)
Year x Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes	No	Yes
Observations	4,590	4,590	4,590	4,590	4,590	4,590
R-squared (panel A)	0.911	0.918	0.701	0.708	0.949	0.951
R-squared (panel B)	0.879	0.896	0.692	0.702	0.900	0.909
Mean DV (panel A)	1.909	1.909	1.259	1.259	3.689	3.689
Mean DV (panel B)	2.665	2.665	1.477	1.477	4.706	4.706

Notes: The unit of analysis is cohort-region-year. Observations weighted by cell size. Dependent variable in the header. Number of hospitalized patients using insurance other than FONASA or ISAPRE (military, other private) in columns 1-2. Patients without insurance and lacking insurance information in columns 3-4 and 5-6. Hospitalization rates adjusted for previous mortality within cell. Population at risk includes all respondents in the 2002 census that reached age 21 between 1964 and 1981 (both years inclusive). Sample period: 2002-2018. "Yr Age 21" is a continuous variable indicating the year when the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C5: Hospitalizations: Type of hospital

	Private hospital		Public hospital: SNSS		Public hospital: SNSS (alternative)		Public hospital: Any Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Female				
Yr Age 21	0.150***		-2.236***		-2.298***		-2.338***	
	(0.026)		(0.099)		(0.101)		(0.100)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	-0.288***	-0.292***	1.338***	0.260**	1.381***	0.261**	1.359***	0.244**
	(0.040)	(0.043)	(0.070)	(0.111)	(0.072)	(0.105)	(0.074)	(0.103)
				Panel B: Male				
Yr Age 21	-0.159***		-3.731***		-3.845***		-3.813***	
	(0.037)		(0.126)		(0.122)		(0.123)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	-0.280***	-0.383***	1.859***	0.702***	1.896***	0.690***	1.742***	0.498***
	(0.047)	(0.053)	(0.070)	(0.171)	(0.071)	(0.156)	(0.070)	(0.156)
Year x Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4,590	4,590	4,590	4,590	4,590	4,590	4,590	4,590
R-squared (panel A)	0.923	0.925	0.884	0.921	0.874	0.915	0.880	0.918
R-squared (panel B)	0.916	0.919	0.908	0.932	0.908	0.935	0.909	0.935
Mean DV (panel A)	16.49	16.49	58.09	58.09	58.98	58.98	60.81	60.81
Mean DV (panel B)	17.35	17.35	57.65	57.65	58.43	58.43	60.73	60.73

Notes: The unit of analysis is cohort-region-year. Observations weighted by cell size. Dependent variable in the header. Number of hospitalized patients at private hospitals in columns 1-2. Patients admitted to public hospitals in the SNSS network in columns 3-4. Columns 5-6 use an alternative definition of SNSS admissions from the discharge summaries. Columns 7-8 use admissions to any public hospital, including those not part of SNSS. Hospitalization rates adjusted for previous mortality within cell. Population at risk includes all respondents in the 2002 census that reached age 21 between 1964 and 1981 (both years inclusive). Sample period: 2002-2018. "Yr Age 21" is a continuous variable indicating the year when the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C6: Mortality Rate at time of discharge of Hospitalized Patients (2004-2012)

	Dependent variable: Deaths at the time of discharge per 1,000 hospitalized patients							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Female				
Yr Age 21	-1.238***							
	(0.092)							
Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)	0.496***	0.141	0.029	0.024	0.060	0.093	0.079	0.104
	(0.120)	(0.188)	(0.163)	(0.161)	(0.158)	(0.164)	(0.152)	(0.177)
				Panel B: Male				
Yr Age 21	-1.502***							
	(0.116)							
Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)	0.625***	0.650***	0.506***	0.419**	0.505**	0.577**	0.550**	0.436
	(0.176)	(0.212)	(0.183)	(0.187)	(0.206)	(0.253)	(0.257)	(0.279)
Year x County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects: Diagnostic...	No	No	Yes	No	No	No	No	No
... x Public insurance...	No	No	No	Yes	No	No	No	No
... x Hospital...	No	No	No	No	Yes	No	No	No
... x Type of admission...	No	No	No	No	No	Yes	No	No
... x Surgery...	No	No	No	No	No	No	Yes	No
... x Previously admitted	No	No	No	No	No	No	No	Yes
Observations (panel A)	603,878	603,878	602,862	601,722	539,443	509,174	493,174	468,553
Observations (panel B)	519,615	519,615	518,681	517,564	455,207	426,039	410,436	390,574
R-squared (panel A)	0.009	0.009	0.180	0.187	0.311	0.345	0.357	0.371
R-squared (panel B)	0.010	0.010	0.174	0.181	0.294	0.319	0.331	0.340
Mean DV (panel A)	11.17	11.17	11.16	11.15	10.59	10.27	10.05	9.892
Mean DV (panel B)	23.56	23.56	23.55	23.55	23.62	23.54	23.28	23.17

Notes: The unit of analysis is a hospitalized patient. Dependent variable is an indicator for whether the patient is dead at time of discharge (multiplied by 1,000). Sample includes patients from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and is limited to one observation per patient (i.e. first admission). Sample period: 2004-2012. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. All regressions include year by county of residence fixed effects. Columns 2-8 replace the baseline cohort trend with age fixed effects. Columns 3-8 iteratively add interactions of fixed effects for diagnostic (column 3: 4-digit ICD code), public insurance (column 4), hospital (column 5: 381 establishments), type of admission (column 6: ER, other establishment, etc.), surgery (column 7), previously admitted (column 8: 2002-2003). Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C7: One-month Mortality Rate of Hospitalized Patients (2004-2012)

	Dependent variable: Deaths within 30 days of discharge per 1,000 hospitalized patients							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Female				
Yr Age 21	-2.119***							
	(0.117)							
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	0.914***	0.307	0.134	0.118	0.125	0.129	0.132	0.140
	(0.165)	(0.230)	(0.189)	(0.187)	(0.200)	(0.214)	(0.207)	(0.210)
				Panel B: Male				
Yr Age 21	-2.713***							
	(0.138)							
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	1.244***	1.066***	0.799***	0.693***	0.726***	0.672**	0.562*	0.448
	(0.205)	(0.222)	(0.213)	(0.214)	(0.252)	(0.296)	(0.311)	(0.340)
Year x County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects: Diagnostic...	No	No	Yes	No	No	No	No	No
... x Public insurance...	No	No	No	Yes	No	No	No	No
... x Hospital...	No	No	No	No	Yes	No	No	No
... x Type of admission...	No	No	No	No	No	Yes	No	No
... x Surgery...	No	No	No	No	No	No	Yes	No
... x Previously admitted	No	No	No	No	No	No	No	Yes
Observations (panel A)	603,878	603,878	602,862	601,722	539,443	509,174	493,174	468,553
Observations (panel B)	519,615	519,615	518,681	517,564	455,207	426,039	410,436	390,574
R-squared (panel A)	0.011	0.011	0.180	0.189	0.305	0.340	0.355	0.370
R-squared (panel B)	0.012	0.012	0.188	0.197	0.304	0.331	0.344	0.352
Mean DV (panel A)	18.83	18.83	18.81	18.81	17.52	16.93	16.58	16.26
Mean DV (panel B)	37.33	37.33	37.31	37.30	36.75	36.33	36.03	35.68

Notes: The unit of analysis is a hospitalized patient. Dependent variable is an indicator for whether the patient dies within 30 days of discharge (multiplied by 1,000). Sample includes patients from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and is limited to one observation per patient (i.e. first admission). Sample period: 2004-2012. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. All regressions include year by county of residence fixed effects. Columns 2-8 replace the baseline cohort trend with age fixed effects. Columns 3-8 iteratively add interactions of fixed effects for diagnostic (column 3: 4-digit ICD code), public insurance (column 4), hospital (column 5: 381 establishments), type of admission (column 6: ER, other establishment, etc.), surgery (column 7), previously admitted (column 8: 2002-2003). Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C8: Six-month Mortality Rate of Hospitalized Patients (2004-2012)

	Dependent variable: Deaths within 6 months of discharge per 1,000 hospitalized patients							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Female				
Yr Age 21	-4.193***							
	(0.188)							
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	2.054***	1.055***	0.767***	0.715**	0.585**	0.631**	0.612**	0.515*
	(0.221)	(0.344)	(0.272)	(0.276)	(0.279)	(0.284)	(0.279)	(0.300)
				Panel B: Male				
Yr Age 21	-5.043***							
	(0.186)							
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	2.402***	1.884***	1.425***	1.287***	1.375***	1.282***	1.189***	1.052***
	(0.255)	(0.328)	(0.334)	(0.334)	(0.380)	(0.360)	(0.374)	(0.391)
Year x County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects: Diagnostic...	No	No	Yes	No	No	No	No	No
... x Public insurance...	No	No	No	Yes	No	No	No	No
... x Hospital...	No	No	No	No	Yes	No	No	No
... x Type of admission...	No	No	No	No	No	Yes	No	No
... x Surgery...	No	No	No	No	No	No	Yes	No
... x Previously admitted	No	No	No	No	No	No	No	Yes
Observations (panel A)	603,878	603,878	602,862	601,722	539,443	509,174	493,174	468,553
Observations (panel B)	519,615	519,615	518,681	517,564	455,207	426,039	410,436	390,574
R-squared (panel A)	0.016	0.016	0.223	0.233	0.336	0.368	0.383	0.397
R-squared (panel B)	0.017	0.017	0.234	0.246	0.346	0.374	0.387	0.396
Mean DV (panel A)	35.46	35.46	35.44	35.43	33.36	32.19	31.42	30.75
Mean DV (panel B)	62.09	62.09	62.09	62.10	60.74	59.72	59.16	58.47

Notes: The unit of analysis is a hospitalized patient. Dependent variable is an indicator for whether the patient dies within six months of discharge (multiplied by 1,000). Sample includes patients from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and is limited to one observation per patient (i.e. first admission). Sample period: 2004-2012. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. All regressions include year by county of residence fixed effects. Columns 2-8 replace the baseline cohort trend with age fixed effects. Columns 3-8 iteratively add interactions of fixed effects for diagnostic (column 3: 4-digit ICD code), public insurance (column 4), hospital (column 5: 381 establishments), type of admission (column 6: ER, other establishment, etc.), surgery (column 7), previously admitted (column 8: 2002-2003). Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C9: Two-year Mortality Rate of Hospitalized Patients (2004-2012)

	Dependent variable: Deaths within two years of discharge per 1,000 hospitalized patients							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Female				
Yr Age 21	-6.872***							
	(0.231)							
Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)	3.438***	2.451***	1.884***	1.808***	1.776***	1.724***	1.716***	1.535***
	(0.265)	(0.430)	(0.344)	(0.355)	(0.360)	(0.388)	(0.370)	(0.400)
				Panel B: Male				
Yr Age 21	-8.137***							
	(0.251)							
Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)	3.739***	3.177***	2.485***	2.282***	2.061***	1.954***	1.775***	1.704***
	(0.345)	(0.413)	(0.397)	(0.388)	(0.445)	(0.475)	(0.483)	(0.489)
Year x County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects: Diagnostic...	No	No	Yes	No	No	No	No	No
... x Public insurance...	No	No	No	Yes	No	No	No	No
... x Hospital...	No	No	No	No	Yes	No	No	No
... x Type of admission...	No	No	No	No	No	Yes	No	No
... x Surgery...	No	No	No	No	No	No	Yes	No
... x Previously admitted	No	No	No	No	No	No	No	Yes
Observations (panel A)	603,878	603,878	602,862	601,722	539,443	509,174	493,174	468,553
Observations (panel B)	519,615	519,615	518,681	517,564	455,207	426,039	410,436	390,574
R-squared (panel A)	0.022	0.022	0.245	0.254	0.345	0.372	0.384	0.398
R-squared (panel B)	0.023	0.023	0.261	0.273	0.368	0.392	0.403	0.414
Mean DV (panel A)	60.04	60.04	60.01	60.01	57.17	55.45	54.38	53.19
Mean DV (panel B)	99.41	99.41	99.39	99.38	97.36	95.82	95.12	93.73

Notes: The unit of analysis is a hospitalized patient. Dependent variable is an indicator for whether the patient dies within two years of discharge (multiplied by 1,000). Sample includes patients from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and is limited to one observation per patient (i.e. first admission). Sample period: 2004-2012. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. All regressions include year by county of residence fixed effects. Columns 2-8 replace the baseline cohort trend with age fixed effects. Columns 3-8 iteratively add interactions of fixed effects for diagnostic (column 3: 4-digit ICD code), public insurance (column 4), hospital (column 5: 381 establishments), type of admission (column 6: ER, other establishment, etc.), surgery (column 7), previously admitted (column 8: 2002-2003). Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C10: Five-year Mortality Rate of Hospitalized Patients (2004-2012)

	Dependent variable: Deaths within five years of discharge per 1,000 hospitalized patients							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Female							
Yr Age 21	-10.566***							
	(0.308)							
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	5.234***	4.180***	3.241***	3.125***	3.061***	2.764***	2.743***	2.424***
	(0.340)	(0.501)	(0.407)	(0.420)	(0.472)	(0.481)	(0.464)	(0.458)
	Panel B: Male							
Yr Age 21	-12.385***							
	(0.327)							
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	5.720***	4.986***	4.090***	3.707***	3.203***	3.052***	2.867***	2.640***
	(0.417)	(0.481)	(0.453)	(0.430)	(0.492)	(0.533)	(0.525)	(0.543)
Year x County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects: Diagnostic...	No	No	Yes	No	No	No	No	No
... x Public insurance...	No	No	No	Yes	No	No	No	No
... x Hospital...	No	No	No	No	Yes	No	No	No
... x Type of admission...	No	No	No	No	No	Yes	No	No
... x Surgery...	No	No	No	No	No	No	Yes	No
... x Previously admitted	No	No	No	No	No	No	No	Yes
Observations (panel A)	603,878	603,878	602,862	601,722	539,443	509,174	493,174	468,553
Observations (panel B)	520,782	520,782	519,844	518,740	456,348	427,135	411,484	391,613
R-squared (panel A)	0.030	0.030	0.238	0.248	0.333	0.358	0.370	0.383
R-squared (panel B)	0.031	0.031	0.247	0.260	0.353	0.377	0.387	0.400
Mean DV (panel A)	92.56	92.56	92.52	92.51	89.05	86.90	85.65	83.82
Mean DV (panel B)	150.1	150.1	150.1	150.1	147.9	146	145.2	142.9

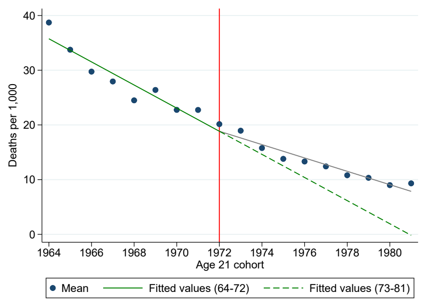
Notes: The unit of analysis is a hospitalized patient. Dependent variable is an indicator for whether the patient dies within five years of discharge (multiplied by 1,000). Sample includes patients from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and is limited to one observation per patient (i.e. first admission). Sample period: 2004-2012. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. All regressions include year by county of residence fixed effects. Columns 2-8 replace the baseline cohort trend with age fixed effects. Columns 3-8 iteratively add interactions of fixed effects for diagnostic (column 3: 4-digit ICD code), public insurance (column 4), hospital (column 5: 381 establishments), type of admission (column 6: ER, other establishment, etc.), surgery (column 7), previously admitted (column 8: 2002-2003). Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C11: Mortality Rate: Unrestricted Sample (2004-2012)

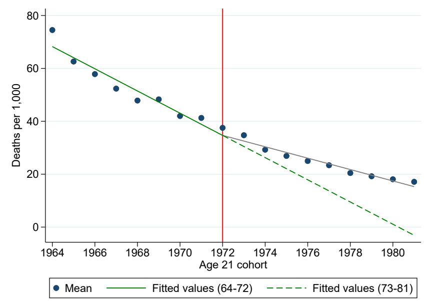
	Dependent variable: Deaths per 1,000 people			
	Female		Male	
	Trend	Age FE	Trend	Age FE
	(1)	(2)	(3)	(4)
	Panel A: Reduced Form			
Yr Age 21	-0.716***		-1.242***	
	(0.025)		(0.039)	
Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)	0.429***	0.076***	0.783***	0.283***
	(0.021)	(0.027)	(0.031)	(0.040)
	Panel B: IV			
Share with college per 10 people	-3.305***	-0.592***	-5.927***	-2.094***
	(0.175)	(0.215)	(0.278)	(0.297)
	Panel C: OLS			
Share with college per 10 people	-2.064***	0.033	-3.601***	-0.433*
	(0.112)	(0.121)	(0.175)	(0.223)
Year x Region FE	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes
Observations	2,430	2,430	2,430	2,430
R-squared (panel A)	0.829	0.864	0.871	0.900
R-squared (panel C)	0.813	0.863	0.849	0.897
Mean DV	4.652	4.652	8.222	8.222
Kleibergen-Paap F-statistic (panel B)	2734	1487	2936	1476

Notes: Sample includes all respondents in the 2002 census that reached age 21 between 1964 and 1981 (both years inclusive), but imposes no restriction on educational attainment. Sample period: 2004-2012. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. Mortality rate and share with college adjusted for previous mortality. Standard errors clustered by region-year in parentheses. *** p<0.01, ** p<0.05, * p<0.1

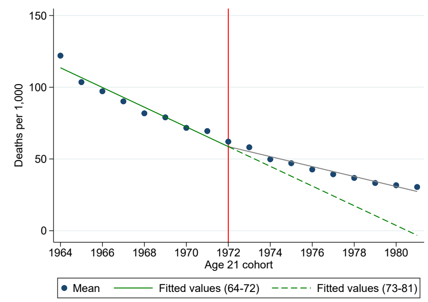
Figure C1: Visualization of Kink: Mortality Rate of Hospitalized Patients (per 1,000)



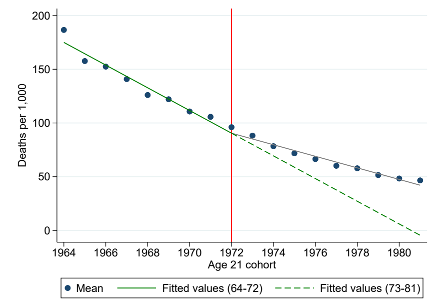
(a) Female: 30 days



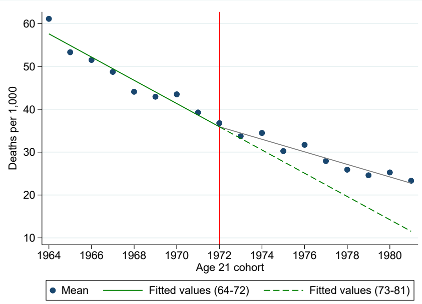
(b) Female: 6 months



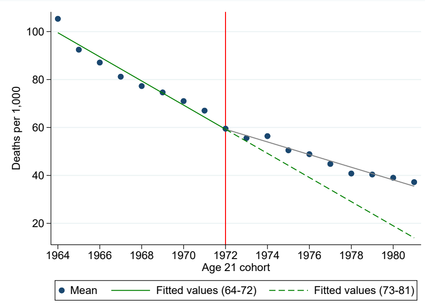
(c) Female: 2 years



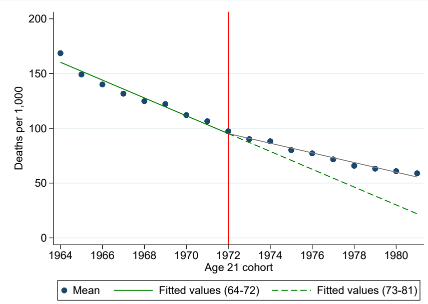
(d) Female: 5 years



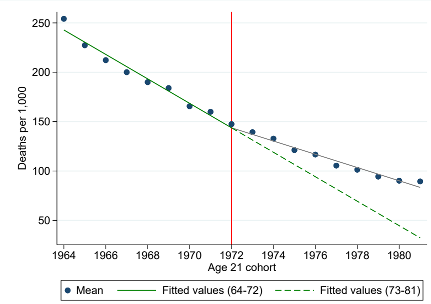
(e) Male: 30 days



(f) Male: 6 months



(g) Male: 2 years



(h) Male: 5 years

Note: Panels show the average mortality rate of hospitalized patients by cohort (averaged across individuals) over the time horizon in the caption (relative to discharge). Sample includes individuals reaching age 21 between 1964 and 1981 (both inclusive), but places no restriction on educational attainment. Sample period: 2004-2012. Solid green line corresponds to the line of best fit for cohorts reaching college age before 1973, which we extrapolate for later cohorts (dashed line). Grey line corresponds to line of best fit for cohorts reaching college age in 1973 or afterwards.

Appendix D Additional Results: CASEN survey

Table D1: Health Insurance and Access to Health Care (Female)

	Enrolled in public health insurance		Received medical care in the last 3 months: General practitioner		Emergency room		Specialist	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Reduced Form				
Yr Age 21	-0.007***		-0.004***		-0.001**		-0.002***	
	(0.001)		(0.001)		(0.001)		(0.001)	
Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)	0.007***	0.005***	0.000	-0.000	0.001	0.001	-0.002*	-0.003**
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
				Panel B: IV				
𝟙(Any College)	-0.269***	-0.200***	-0.011	0.009	-0.024	-0.027	0.076*	0.101**
	(0.053)	(0.062)	(0.066)	(0.070)	(0.030)	(0.029)	(0.043)	(0.046)
				Panel C: OLS				
𝟙(Any College)	-0.240***	-0.240***	-0.002	-0.002	-0.013***	-0.013***	0.031***	0.031***
	(0.007)	(0.007)	(0.004)	(0.004)	(0.002)	(0.002)	(0.004)	(0.004)
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	81,893	81,893	57,601	57,601	78,822	78,822	82,801	82,801
R-squared (panel A)	0.185	0.187	0.080	0.081	0.066	0.067	0.084	0.085
R-squared (panel C)	0.225	0.228	0.080	0.081	0.067	0.068	0.085	0.086
Mean DV	0.627	0.627	0.220	0.220	0.0665	0.0665	0.145	0.145
Kleibergen-Paap F-statistic (panel B)	291.3	209.2	213.1	149	280.6	197.6	293	214

Notes: The unit of analysis is a female respondent from the CASEN survey. Dependent variable in the header. A dummy for being enrolled in the public health insurance (FONASA) in columns 1-2; a dummy for having seen a general practitioner in the past three months in columns 3-4; similar dummies for visits to the emergency room or to a specialist in columns 5-6 and 7-8 respectively. Sample period: 1990-2017 (13 waves). Not all questions are asked in all years. Sample includes individuals from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and report 4+ years of secondary education. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. All regressions include year by county of residence fixed effects. Even-numbered columns replace the baseline cohort trend with age fixed effects. Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C2: Health Insurance and Access to Health Care (Male)

	Enrolled in public health insurance		Received medical care in the last 3 months: General practitioner		Emergency room		Specialist	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Reduced Form				
Yr Age 21	-0.009***		-0.003***		-0.001***		-0.004***	
	(0.001)		(0.001)		(0.000)		(0.001)	
Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)	0.012***	0.009***	-0.003**	-0.005***	0.001	-0.000	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
				Panel B: IV				
𝟙(Any College)	-0.528***	-0.410***	0.119**	0.217***	-0.026	0.013	-0.014	0.052
	(0.065)	(0.068)	(0.057)	(0.056)	(0.023)	(0.028)	(0.039)	(0.042)
				Panel C: OLS				
𝟙(Any College)	-0.234***	-0.233***	0.015***	0.015***	-0.010***	-0.010***	0.031***	0.032***
	(0.006)	(0.006)	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.003)
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	79,580	79,580	55,366	55,366	76,682	76,682	80,693	80,693
R-squared (panel A)	0.173	0.175	0.081	0.082	0.060	0.061	0.081	0.083
R-squared (panel C)	0.212	0.213	0.081	0.082	0.060	0.061	0.083	0.085
Mean DV	0.555	0.555	0.153	0.153	0.0447	0.0447	0.0877	0.0877
Kleibergen-Paap F-statistic (panel B)	195.9	181.5	123.3	108.7	176.2	176.6	196.5	183.7

Notes: The unit of analysis is a male respondent from the CASEN survey. Dependent variable in the header. A dummy for being enrolled in the public health insurance (FONASA) in columns 1-2; a dummy for having seen a general practitioner in the past three months in columns 3-4; similar dummies for visits to the emergency room or to a specialist in columns 5-6 and 7-8 respectively. Sample period: 1990-2017 (13 waves). Not all questions are asked in all years. Sample includes individuals from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and report 4+ years of secondary education. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x 𝟙(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. All regressions include year by county of residence fixed effects. Even-numbered columns replace the baseline cohort trend with age fixed effects. Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C3: Access to Health Care: Additional results

	Received medical care in the last 3 months:									
	Preventive care		Sickness		Exams		Hospital/surgery		Dental	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					Panel A: Reduced Form					
Yr Age 21	-0.004***		-0.001**		-0.005***		-0.001***		0.000	
	(0.001)		(0.001)		(0.001)		(0.000)		(0.000)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	0.001	-0.003***	0.001	-0.001	0.000	-0.002***	-0.001	-0.002***	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
					Panel B: IV					
$\mathbb{1}$(Any College)	-0.031	0.141***	-0.023	0.047	-0.019	0.091***	0.022	0.066***	0.042	0.026
	(0.031)	(0.039)	(0.035)	(0.053)	(0.029)	(0.031)	(0.015)	(0.017)	(0.026)	(0.030)
					Panel C: OLS					
$\mathbb{1}$(Any College)	-0.004	-0.003	0.006***	0.006***	0.018***	0.018***	0.004**	0.004***	0.046***	0.046***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
County x Year x Gender FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	76,917	76,917	50,469	50,469	163,503	163,503	163,101	163,101	163,504	163,504
R-squared (panel A)	0.101	0.103	0.740	0.740	0.149	0.151	0.750	0.750	0.070	0.071
R-squared (panel C)	0.101	0.103	0.740	0.740	0.150	0.151	0.750	0.750	0.074	0.074
Mean DV	0.086	0.086	0.247	0.247	0.176	0.176	0.202	0.202	0.110	0.110
Kleibergen-Paap F-statistic (panel B)	323.3	231.4	201.3	75.42	378.9	341.6	382	347.9	378.6	341.5

Notes: The unit of analysis is an individual respondent from the CASEN survey. Dependent variable in the header. A dummy for having visited a preventive care physician in the past three months in columns 1-2; similar dummies for visits due to sickness or accident in columns 3-4; medical exams or tests in columns 5-6; hospitalization or surgery in columns 7-8 and dentist in columns 9-10. Sample period: 1990-2017 (13 waves). Not all questions are asked in all years. Sample includes individuals from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and report 4+ years of secondary education. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. All regressions include year by county of residence by gender fixed effects. Even-numbered columns replace the baseline cohort trend with age fixed effects. Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C4: Health Behaviors and Status (Female)

	Smoked in last month		Pap smear in last 3 years		Self-assessment: good health		Sick or in accident in last 3 months	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Reduced Form				
Yr Age 21	0.015***		0.019***		0.006***		-0.004***	
	(0.002)		(0.002)		(0.001)		(0.001)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)	-0.013***	-0.009**	-0.012***	-0.004**	0.002	0.002	0.001	0.000
	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
				Panel B: IV				
$\mathbb{1}$(Any College)	0.460***	0.332**	0.498***	0.168**	-0.070	-0.068	-0.047	-0.019
	(0.126)	(0.163)	(0.095)	(0.084)	(0.088)	(0.089)	(0.047)	(0.049)
				Panel C: OLS				
$\mathbb{1}$(Any College)	0.016	0.014	0.021***	0.020***	0.100***	0.100***	0.003	0.003
	(0.010)	(0.009)	(0.005)	(0.005)	(0.007)	(0.007)	(0.003)	(0.003)
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	14,774	14,774	58,549	58,549	46,098	46,098	82,241	82,241
R-squared (panel A)	0.053	0.056	0.103	0.111	0.098	0.099	0.100	0.100
R-squared (panel C)	0.052	0.056	0.103	0.111	0.104	0.105	0.100	0.100
Mean DV	0.359	0.359	0.726	0.726	0.497	0.497	0.211	0.211
Kleibergen-Paap F-statistic (panel B)	68.27	43.31	214.2	131.8	217.3	147.1	290.1	214

Notes: The unit of analysis is a female respondent from the CASEN survey. Dependent variable in the header. A dummy for smoking in the past month in columns 1-2; a dummy for having a Pap smear in last three years in columns 3-4; a dummy for self-assessed health status good or very good (6-7 on 7-point scale or 4-5 on 5-point scale) in columns 6-7; a dummy for being sick or involved in an accident in the past 3 months in columns 7-8. Sample period: 1990-2017 (13 waves). Not all questions are asked in all years. Sample includes individuals from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and report 4+ years of secondary education. “Yr Age 21” is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. “Yr Age 21 x $\mathbb{1}$(Yr Age 21 ≥ 1973)” is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. All regressions include year by county of residence fixed effects. Even-numbered columns replace the baseline cohort trend with age fixed effects. Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C5: Health Behaviors and Status (Male)

	Smoked in last month		Self-assessment: good health		Sick or in accident in last 3 months	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Reduced Form			
Yr Age 21	0.015***		0.006***		-0.002***	
	(0.003)		(0.001)		(0.001)	
Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)	-0.015***	-0.010**	0.004**	0.003	-0.001	-0.002
	(0.004)	(0.004)	(0.002)	(0.002)	(0.001)	(0.001)
			Panel B: IV			
$\mathbb{1}$(Any College)	0.768***	0.529**	-0.224**	-0.149	0.027	0.092
	(0.238)	(0.255)	(0.100)	(0.116)	(0.047)	(0.058)
			Panel C: OLS			
$\mathbb{1}$(Any College)	-0.038***	-0.040***	0.093***	0.094***	0.009***	0.009***
	(0.010)	(0.010)	(0.006)	(0.006)	(0.003)	(0.003)
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	Yes	No	Yes	No	Yes
Observations	14,839	14,839	40,608	40,608	80,175	80,175
R-squared (panel A)	0.051	0.053	0.109	0.110	0.090	0.091
R-squared (panel C)	0.051	0.054	0.115	0.116	0.090	0.091
Mean DV	0.435	0.435	0.558	0.558	0.147	0.147
Kleibergen-Paap F-statistic (panel B)	31.05	16.03	69.83	59.23	194.7	181.6

Notes: The unit of analysis is a male respondent from the CASEN survey. Dependent variable in the header. A dummy for smoking in the past month in columns 1-2; a dummy for self-assessed health status good or very good (6-7 on 7-point scale or 4-5 on 5-point scale) in columns 3-4; a dummy for being sick or involved in an accident in the past 3 months in columns 5-6. Sample period: 1990-2017 (13 waves). Not all questions are asked in all years. Sample includes individuals from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and report 4+ years of secondary education. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. In panel B, this variable is used as excluded instrument for the share with college. All regressions include year by county of residence fixed effects. Even-numbered columns replace the baseline cohort trend with age fixed effects. Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C6: Reasons for not Having a Pap Smear

	Pap smear test in last 3 years					
		No + stated reason				
	Yes	Doesn't know about it	Doesn't know where to do it	Afraid or doesn't like it	Forgot to do it	Uninterested or doesn't need it
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Any College)	0.168**	-0.016*	-0.001	-0.009	-0.075	-0.014
	(0.084)	(0.009)	(0.007)	(0.033)	(0.049)	(0.048)
County x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	58,549	58,549	58,549	58,549	58,549	58,549
Mean DV	0.726	0.00316	0.00162	0.0329	0.0960	0.0958
Kleibergen-Paap F-statistic	131.8	131.8	131.8	131.8	131.8	131.8

Notes: The unit of analysis is an individual female respondent from the CASEN survey. Dependent variable in the header. In column 1, a dummy for having had a Pap smear in the past three years. In columns 2-6 respective dummies for stated causes for not having a Pap smear (set to zero if Pap smear in past three years). These causes are not exhaustive. Survey waves: 1992, 1994, 1996, 2000, 2003, 2009, 2011, 2013, 2015, 2017. Sample includes individuals from cohorts that reached age 21 between 1964 and 1981 (both years inclusive) and report 4+ years of secondary education. "Yr Age 21" is a continuous variable indicating the year at which the cohort reached age 21, normalized to zero in 1972. "Yr Age 21 x $\mathbb{1}$(Yr Age 21 $\geq$ 1973)" is the interaction of this variable with a dummy for cohorts that reached age 21 on or after 1973. We use this variable as excluded instrument for the share with college. All regressions include year by county of residence and age fixed effects. Standard errors clustered two-way by county and region-year reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.