No Calm Before the Storm: Economic Losses Prior to Dementia Diagnoses among Working-Age Colombians*

Leonardo Bonilla Mejía
Banco de la República

María Fernanda Henao Pineda
Universidad de Manizales

Nicholas W. Papageorge†
Johns Hopkins University and NBER

Preliminary and Incomplete.

Abstract: Much research examines the consequences of dementia due to Alzheimer's Disease and Related Dementias (ADRD) among the elderly (age 65+). ADRD rarely affects younger people (between ages 30-64). Yet, other forms of dementia do. Many are due to trauma or risky behavior. We know very little about the social, health, and economic correlations and implications of dementia among younger people, and thus lack a full picture of the population implications of dementia. This paper examines all forms of dementia diagnoses among Colombians across the age distribution, with particular focus on working-age Colombians. We present three main findings. One, differences between people with a dementia diagnosis and those without tend to emerge 5-10 years prior to diagnosis for many outcomes, including labor supply and wages. Second, impacts of dementia on younger groups extend to novel outcomes, such as criminality, in line with the idea that dementia among younger people suggests the need to focus on outcomes more likely to be associated with non-elderly individuals. Third, mental health is a key comorbidity of dementia, driven almost entirely by addiction disorders. These preliminary results suggest that an examination of younger individuals permits a more complete accounting of dementia. Moreover, our results are in line with a growing literature showing that dementia affects more people on more dimensions and at younger ages than has generally been understood, again suggesting a need to look at younger people to fully appreciate the consequences of dementia.

Keywords: Alzheimer's Disease, Dementia, Labor, Credit, Crime, Aging.
JEL Classification: J14, I12, I14, J22, J26, D14, G51, G52

*First version: August 4, 2025: Current version: March 16, 2026. We gratefully acknowledge helpful comments from Jon Skinner and Kevin Thom, as well as from seminar participants at Banco de la República and EAFIT, and members of the Grupo de Neurociencias de Antioquia (GNA). The authors would also like to thank the excellent research assistance of María José Otero. The opinions contained in this document are the sole responsibility of the authors, and do not commit the Banco de la República or its Board of Directors.

†Corresponding author: papageorge@jhu.edu

1 Introduction

Reams of research have shown that dementia constitutes a massive and costly health and economic challenge. Research tends to focus on Alzheimer's Disease and Related Dementias (hereinafter *ADRD*), which are "neurodegenerative forms (herein *ND*) of dementia that nearly exclusively affect the elderly and are generally neither treatable nor reversible. Beyond medical expenditures, lost productivity, early mortality, and patient suffering, recent papers have identified multitudes of indirect consequences of ADRD (e.g., on labor supply, retirement, portfolio decisions, and housing choices ((Li et al., 2022, 2023; Jeong et al., 2025; Mazzonna and Peracchi, 2024)), along with costly spillover effects (e.g., on family members, especially labor supply of daughters who often serve as informal caregivers (Shen, 2024)).[1]

It is well-documented that ADRD is a growing problem due to aging populations. Moreover, it likely already represents an even broader challenge than is generally understood. One reason is under-diagnosis, which is estimated to be massive: roughly 2/3 of cases in the U.S. are not diagnosed (Lang et al., 2017). Another reason is that recent studies have increasingly shown that the adverse impacts of ADRD begin to emerge long before diagnosis, on average about 10 years, but up to 20 years before (Li et al., 2022; Nicholas et al., 2021). Recent evidence using genetic risk of ADRD also suggest that economic consequences of ADRD can begin long before diagnosis and, moreover, often extend to many individuals who are never diagnosed with ADRD (Jeong et al., 2025).

Most ADRD research is focused on the elderly as the condition is generally exceedingly rare among younger populations, though there are some exceptions, including a genetic mutation unique to Colombia that leads to early-onset ADRD.[2] Never-the-less, other forms of dementia exist that do affect younger populations. The presentation is often similar to ADRD (e.g., memory loss and diminished cognitive function), but these forms of dementia are usually "non-neurogenerative" (henceforth *non-ND*) which means they are often attributable to environmental factors (e.g., substance abuse or trauma due to accidents) and in some cases are reservable and treatable, which stands in stark contrast to ADRD.

Given that the focus of dementia research is on the elderly, we know relatively little about the extent and consequences of dementia among younger, working age populations and thus

[1] See Chandra et al. (2023) for an overview of research from economics on Alzheimer's Disease.

[2] The so-called *Mutación Paisa* or *Paisa Mutation*) leads to early-onset dementia. "Paisa" is a common term referring to people in and near a municipality of Colombia, Antioquia, of which the city Medellín is the capital. The Paisa Mutation is most prevalent in this region. For an overview of the Paisa Mutation, including ongoing work related to it, see, for example, Acosta-Baena et al. (2011) and Aguirre-Acevedo et al. (2016) along with other work performed by the institute responsible for discovering the genetic mutation, the "Grupo de Neurociencias de Antioquia". While this paper is not about the Paisa Mutation, it is related in the sense that it aligns with the idea of considering younger populations when examining dementia.

lack a complete picture of the full consequences of dementia in the population. Indeeed, in many contexts, few younger people are screened for dementia since the condition is largely viewed as irrelevant for them (Blue Cross Blue Shield Association, 2020). Consistent with this view, longitudinal data sets do not facilitate investigation of younger or "early onset" dementia (defined as dementia diagnosed between ages 30-64). Data sets that do include rich information pertinent to dementia, such as the Health and Retirement Study in the U.S., are focused on people over age 50 and tend to have less information about earlier life outcomes. Similarly, data sets focusing on younger people (e.g., the Survey of Adolescent Health) tend to not include information on dementia or measures of cognitive decline (Harris and Halpern, 2022). Given recent research showing that the consequences of ADRD occur earlier than previously thought and given that non-ND can affect people across the age distribution, a full accounting of dementia would require an examination that is not limited solely to the elderly.

This paper studies the health and economic trajectories of people with dementia diagnoses in Colombia. While our analysis includes pre- and post-diagnosis trends for all diagnosed individuals, we place explicit focus on pre-diagnosis trends in health and economic outcomes and on individuals diagnosed when relatively young (between ages 40-64). As we explain in more detail below, in our main analyses, we compare diagnosed individuals to control groups we construct using observably similar individuals who are never diagnosed with any form of dementia in the years prior to diagnosis. Our data allow us to distinguish between people with ND dementia (due to ADRD) along with non-ND dementia, which is more common among younger people.

Our focus on younger people helps to motivate our analyses. Because dementia is generally seen as a problem among the elderly, the focus of medical and social scientific research on dementia is often on outcomes related to the elderly, such as retirement age or pension wealth. However, if individuals begin suffering earlier (in line with recent work on ADRD or due to non-ND), it is important to understand how dementia links to outcomes across the lifecycle (Li et al., 2022; Nicholas et al., 2021). To that end, we focus on how diagnosis relates to health trajectories, labor market behavior and outcomes, credit use, and criminality among individuals from different age groups in the years leading up to diagnosis. Exploring how criminality and dementia relate is a novel feature of our study and arises directly from the idea that a full accounting of dementia naturally leads to an examination of youth-related outcomes.

To our knowledge, this is the first paper that considers dementia and economic links in the Colombian context and the first paper to study lifecycle outcomes of large populations diagnosed with early-onset dementia in general. We focus on the Colombian context for

several reasons. First, Colombian health and labor data are extensive and linkable, and include long histories since individuals are followed using their national identification number (see Bernal et al. (2017); Camacho and Conover (2013); Buitrago et al. (2021) for other papers using these data). This facilitates an examination of people years prior to diagnosis. Second, Colombia has robust healthcare and pension systems, including disability pensions, which covers thorough ADRD diagnoses (MINSALUD, 2020). Third, given the presence of the aforementioned genetic mutation leading to early-onset ADRD, there is some evidence that younger people are more likely to be screened for dementia in Colombia compared to other countries. Fourth, we are interested in how dementia relates to economic behavior and health outcomes in an emerging economy as dementia is not only a first-world problem.

Our sample is based on administrative healthcare utilization data from Colombia, known as the *Registros Individuales de Prestación de Servicios de Salud* (RIPS), which we match to formal employment and earnings data from the *Planilla Integrada de Liquidación de Aportes* (PILA). While the RIPS data set captures detailed diagnostic information, it only includes individuals who access formal health services; The PILA data set, in turn, excludes informal workers. We also incorporate demographic and diagnostic data to identify Alzheimer's disease and related dementias. Health data from RIPS includes over 894,723 diagnosed individuals, of whom 253,534 can be linked to formal labor records. For our main analysis, we focus on first diagnoses between 2020 and 2022, covering 143,967 individuals in the health data. Of these, 44,870 individuals appear in the labor data set and thus comprise our main analysis sample. To ensure diagnostic quality, we require specialist visits, appropriate diagnostic testing, and exclude emergency room diagnoses, following Colombia's clinical practice guidelines and expert consultation. These restrictions yield 70,261 individuals in our main health analysis sample, and 22,565 in the linked labor data. We supplement these data with credit records from the financial regulator and criminal records from the judicial branch to examine financial and legal outcomes.

Our descriptive analysis shows several intriguing patterns. First, annual diagnoses fluctuate quite a bit between about 2000-4000 per year but that there is no clear evidence of trends over time, i.e., the average is roughly flat, which helps to dispel concerns that any subsequent patterns are due to changes in how diagnosis occurs. Second, there is clear evidence that dementia across the age distribution. As we would expect, most dementia occurs among older people (65 and up) but about 17% of diagnosis occur among people younger than 60. About 5% are among people younger than 40. Third, and also expected, nearly all dementia among the very young is non-ND. Fourth, while average numbers of diagnoses do not show any obvious trend over time, there are some differences in age patterns, in particular, more youth diagnoses in later versus earlier years (e.g., 2009 versus 2020). In part, this pattern

may reflect growing understanding of the Paisa mutation. Fifth, women with dementia diagnoses outnumber men (about 43,000 versus 27,000) and tend to be slightly older. Finally, given our sample restrictions, we do not seem to be capturing regional patterns reflecting genetic mutations; if we compare the area with the aforementioned Paisa mutation and other parts of Colombia, we find that the age distribution is fairly similar.[3]

Our main analysis aims to understand dementia in the years leading up to diagnosis. Our principle method is to construct a diagnosed group and a control group of never-diagnosed individuals, matched on age, sex, and municipality. We compare trajectories over time in the years prior to a diagnosis, both by plotting coefficients and then as a formal event study. We focus on the pre-diagnosis period and on heterogeneity across gender and age to measure the pre-diagnosis costs of illness. For each outcome of interest, we consider all dementias followed by ADRD and no-ND. For each of these three groups, we consider all individuals, and then separately consider those aged 40-60 and those age 65 or more.Additionally, we consider even younger individuals, though sample sizes are small and results thus noisey. Still, for completeness, we include them in the Appendix.

We begin our analysis with health care utilization (measured as log expenditures). We provide clear evidence of higher utilization among people who are eventually diagnosed. On average, substantial differences begin to emerge about 2 years prior to diagnosis. Patterns are strikingly similar for ADRD versus non-ND and for different age groups, though gaps between eventually diagnosed and the never diagnosed control group tend to emerge earlier for those diagnosed at ages 40-64: about 3-4 years prior to diagnosis. Similar patterns are found if we consider health events, e.g., doctors' visits or procedures, instead of expenditures.

Our next set of analyses aims to shed light on what types of medical needs and commorbidities might be driving increased healthcare utilitzation. We use "co-occurrence networks", a method that allows one to analyze diagnoses that tend to co-occur (Dervić et al., 2025); (Jensen et al., 2014). For people with dementia, this analysis suggests that circulatory problems, metabolic conditions, and addictions appear to co-occur with frequency among the eventually diagnosed, but not among the control group. In particular, among controls, we also find co-occurence of circulatory and metabolic health problems, consistent with an aging population. Among diagnosed we also see co-occurence of of circulatory and metabolic health conditions along with mental health problems. Further analysis shows that the most substantial source of mental health problems for this group is some form of addiction. In other words, a distinguishing feature of people who are eventually diagnosed with dementia

[3]If we are less strict in sample selection to include potential mistaken disagnoses, we see younger age patterns likely due to higher rates of diagnosis among youth if doctors, especially non-specialists, in the region are more likely to screen for or diagnose dementia.

is a history of mental health problems, in particular, addiction problems. These findings suggest that a precursor of diagnosed dementia is substance abuse, which can lead to non-ND dementia among younger people and seems to correlate to ADRD. The source of the correlation is, however, unclear and would require further analysis. While it is possible that addiction increases ADRD risk, there are other possibilities, e.g., that early stages of ADRD lead to mental health struggles, in turn leading people to self-medicate with addictive substances.

We continue our analysis of dementia across the age distribution by examining labor outcomes. Labor supply gaps between the diagnosed and the control groups are stark and emerge earlier. Among individuals aged 40-64 when first diagnosed, differences emerge 5-7 years prior to diagnosis. For individuals diagnosed after age 65, differences emerge earlier, which likely reflects earlier impacts of dementia that go undiagnosed for some years and affect labor. We find similar patterns for wages and if we focus on men or women separately, though mens' labor outcome differences between diagnosed and controls are more pronounced.

Finally, we conduct a simlar analysis (comparing enventually diagnosed and control groups in the 10 years leading up to diagnosis), considering amount of outstanding credit and conviction of crimes. In both cases, the evidence is somewhat weak due to noisiness driven by small samples. Yet, there is evidence of larger outstanding credit balances and of higher rates of convicion for a crime, especially among the younger group diagnosed (between 40-64). These findings underscore the importance of considering younger populations—and outcomes that are more associated with younger people—to obtain a more complete understanding of the full population implications of dementia.

Our study relates to a literature on Alzheimer's Disease and related dementias that considers the indirect or spillover effects on economic outcomes.(Nicholas et al., 2021) quite similar to our paper in comparing never- to eventually-diagnosed individuals and finding large effects many years before diagnosis. (Li et al., 2022) find similar effects. Many of these papers rely on diagnosed individuals as the treatment group, which means individuals who may be afflicted by ADRD may constitute part of the control group. However, (Li et al., 2022) use additional measures of dementia and finds similar gaps. Moreover, an additional set of papers examines the affect not of diagnosis or eventual diagnosis, but instead of genetic risk of ADRD and also finds generally that effects begin not only quite early (e.g., during work age), but also among people who are never diagnosed.[4] The current paper also relies on diagnosis, which means we face issues of selection and must interpret findings accordingly. However, we add to the literature on the indirect costs of ADRD in a unique way by exploring early-onset dementia among a large population with information on a host

[4](Borgbjerg et al., 2024) and (Li et al., 2024) also consider labor outcomes and genetic data for ADRD.

of outcomes relevant to their stage of the lifecycle. To our knowledge, this is the first paper to use administrative data to relate early-onset dementia to lifecycle outcomes in a non-clinical setting.

Our paper also contributes to a large and well-developed literature on the interconnectedness of health and other economic factors. While many papers have made this point in a variety of contexts, our work is more closely aligned to papers studying chronic medical conditions that affect younger individuals and thus relate to factors such as labor supply and earnings. Examples are depression (Zimmerman and Katon, 2005), HIV/AIDS (Thirumurthy et al., 2008), addiction (French et al., 1997), diabetes (Brown et al., 2005) and, of course, COVID-19 (Albanesi and Kim, 2021). Our point is that ADRD may be another condition that affects people's work lives and other economic outcomes, but at younger ages than has previously been understood.

Finally, and more broadly, our work relates to the basic idea of selection into treatment and diagnosis. This is a well-known problem in medicine and in the social sciences that tackle health-related topics. Some papers model selection (Cronin et al., 2020; Papageorge, 2021). Others try to correct for it, e.g., with instruments (Polsky and Basu, 2012). The general view is that it is difficult to get around selection into diagnosis or into medical treatment. While we do not solve the problem here, we do shed light on a possibly severe form of selection into diagnosis where conventional knowledge and current practice lead to severe underdiagnosis. In our case, a dreadful genetic mutation may have silver lining, prompting medical professionals to consider the unthinkable: that young people may be suffering from dementia. Whether the rates of dementia we observe here carry over to other countries is an open question. If so, the type of work we are conducting here relating early-onset dementia to early-lifecycle outcomes will provide important insights into the extent and nature of the condition.

The remainder of this paper is organized as follows. In Section 2, we discuss the data used in this paper, construction of the main analysis sample, and preliminary descriptive statistics. Section 3 introduces our empirical approach. Section 4 presents our results. Section 5 concludes.

2 Data Description and Preliminary Data Analysis

2.1 Main Health and Labor Sample

The primary data source used in our analysis is the *Registro Individual de Prestación de Servicios de Salud* (RIPS), which translates to "Register of Individual Health Services."

The RIPS data set spans the period 2009 to 2022 and compiles mandatory reports from all health service providers nationwide, including detailed information with dates on medical diagnoses, procedures, appointments, and hospitalizations. According to Panel A of Table 1, 49.5 million individuals are recorded in RIPS, of whom 894,723 individuals are diagnosed with any dementias (grouped as ADRD and non-ND). Diagnoses are coded using the International Classification of Diseases, 10th Revision (ICD-10). Of these diagnosis, 592,996 are diagnosed with ADRD and 301,727 are diagnosed with non-ND. We note that Alzheimer's Disease is more difficult to diagnose than other dementias, so many individuals with a dementia diagnosis could be suffering from Alzheimer's. This motivates our consideration of ADRD and non-ND, both separately and together.

To incorporate information on labor market participation, we link the RIPS data to the *Planilla Integrada de Liquidación de Aportes* (PILA), which translates to "Integrated Roster of Benefit Payments". The PILA data report monthly information on formal employment and earnings. This system is used by employers and self-employed workers to declare payroll contributions to health, pension, and social security . The PILA database also includes information on wages, firm characteristics, job tenure, and industry of employment. The second column of Panel A in Table 1 shows that 26,840,329 individuals are observed in both the RIPS and PILA data sets. Of these individuals, 164,110 have an ADRD diagnosis and 89,433 with a non-ND diagnosis.

2.2 Main Sample Construction

To construct our analysis sample, we first focus on individuals with a dementia diagnosis (the "treated" group), and then construct a "control" group of otherwise similar individuals. For the treated group, we restrict attention to the sample of 253,534 individuals who appear in both the healthcare (RIPS) and the labor market (PILA) data sets and who have any dementia diagnosis. We then restrict the sample to individuals for whom we observe a first diagnosis between 2020 and 2022. Sample sizes are found in Panel B of Table 1. We restrict to these years to permit analysis of long trajectories leading up to the data of diagnosis, though it is straightforward to include individuals with earlier diagnoses; indeed, for some descriptive statistics, we include earlier years of first diagnosis, which we describe below. In the restricted sample, (Panel C of Table 1) we observe 143,967 individuals in RIPS and 44,870 in the linked RIPS-PILA data. Among them, 65,380 and 78,587 individuals, respectively, are observed with a first ADRD and non-ND diagnosis during our sample period.

To ensure diagnostic quality and reduce misclassification, we impose additional restric-

tions based on clinical guidelines and expert consultation.[5] First, we require that individuals visited at least one relevant specialty (endocrinology, genetics, geriatrics, internal medicine, neurology, psychology, or psychiatry) in the year prior to diagnosis. Second, we require that individuals received at least one diagnostically relevant medical examination in the year prior to diagnosis, as specified in Colombia's Clinical Practice Guidelines for diagnosis and treatment of major neurocognitive disorder (dementia).[6] Third, we exclude individuals whose first recorded diagnosis occurs during an emergency room visit, as these may represent acute events rather than confirmed dementia diagnoses. These restrictions yield our main analysis sample of 70,261 individuals with clinically validated dementia diagnoses.

To establish the control group, individuals were identified from the PILA and RIPS databases who were active within the same period as the treated group and had no prior diagnosis of any dementia. Controls were matched to treated individuals based on age, sex, and municipality of residence at the time of the treated individual's diagnosis, and based on whether the individual had ever participated in the labor market. Following the identification of potential controls, a **2:1 ratio** was applied, with two controls selected for each treated individual. To ensure fair and uniform assignment, controls within each demographic stratum (age-sex-municipality) were randomly ordered and then assigned cyclically to treated individuals. Each control was assigned a **synthetic diagnosis date** corresponding to that of their matched treated individual. This methodology enables an event study analysis where both groups share a synchronized timeline relative to the time of diagnosis of treated group individuals.

Finally, we note that, although administrative records offer large-scale coverage and detailed information on both healthcare and employment, they come with limitations. RIPS only captures individuals who interact with the formal healthcare system, potentially missing undiagnosed or untreated cases. As in many other contexts, the number of actual all dementia cases in Colombia (as in other parts of the world) is likely to be much higher due to under-diagnosis. Similarly, PILA includes only formal sector workers, excluding those employed in the informal economy. Over half of working Colombians work in the informal sector (DANE, 2024), and they are therefore excluded from our analysis sample even though it is possible they are observed in the RIPS (healthcare) data set. We are unable to distinguish between diagnosed people who do not work versus those who work in the informal sector. Thus, any results are based on a selected sample of individuals who tend to be more attached to the formal labor market and who use formal healthcare. These individuals tend

[5]These criteria were developed in consultation with clinical experts from the Grupo de Neurociencias de Antioquia (GNA), Universidad de Antioquia.

[6]Ministerio de Salud y Protección Social (2013), *Guía de Práctica Clínica para el diagnóstico y tratamiento del trastorno neurocognoscitivo mayor (Demencia)*, Sistema General de Seguridad Social en Salud, Colombia.

to be of higher socioeconomic status compared to the the typical Colombian, but may also be more similar to individuals in wealthier countries where use of formal healthcare and engagement in formal work are much more typical.

We note that removing these filters to include more diagnosed individuals does not change results qualitatively, but tends to lead to larger numbers of younger individuals with a (potentially erroneous) dementia diagnosis. Sample counts are in Panel C of Table 1. Taking into account *Column 3*, 13,736 individuals are found in both RIPS and PILA, have an observed identification number, and were diagnosed with ADRD between 2020-2022 according to the aforementioned criteria. Corresponding numbers for ADRD and non-ND are 5,410 and 8,326, respectively.

2.3 Secondary Data: Credit and Crime

For additional analysis, we also consider credit and crime, which we match to our main sample from RIPS and PILA, constructed as described above.

2.3.1 Credit

To obtain data on credit, wee use loan-level credit data from the *Superintendencia Financiera de Colombia* (SFC) Format 341, (a record of loans from the Finance Ministry) spanning 2009-2022. These mandatory regulatory reports cover all active credit operations in Colombia's financial system at the individual borrower-lender level, including credit issuance dates, loan amounts, interest rates, maturities, and credit type classifications. We link the credit data to healthcare records using national ID numbers. Of the 49,5 million individuals in the healthcare database, approximately 30% have valid national IDs that enable linkage to the credit system. Within our matched treatment sample, we identify 195,730 individuals with any credit market participation during 2009-2022, including 126,475 diagnosed with ADRD and 69,254 with non-ND.

2.3.2 Crime

Criminal records are obtained from Colombia's Ministry of Justice *Rama Judicial Consejo Superior de la Judicatura*, which maintains publicly accessible online records for convicted defendants. We extract criminal histories for a random 33% subsample of 11,280 individuals identified in the credit data. Using national ID numbers, we retrieve the complete criminal history for each individual in our sample during 2009-2022. The records include defendant

identification, crime type, crime date, and sentence information.[7]

2.3.3 Descriptive Empirical Patterns

Table 3 compares treatment and control variables across our four datasets before and after the diagnosis date of the treated individuals, i.e., those who appear with any dementia diagnosis between the years 2020-2022. Pre- and post-diagnosis age and gender ratios are roughly equivalent for the treatment and control. Average levels of formal employment and salary are more similar in the pre-diagnosis period compared to the post-diagnosis period. The table also considers labor supply (extensive margin) and salary for males and females separately. In general, we find that individuals in control groups work more than those treatment groups (again, on the extensive margin), and that there are larger differences for females. Wage differences for control versus treatment groups are also larger for males, but exist among females as well.

Figure 1 shows diagnosis of ADRD and non-ND across the sample period. We observe substantial numbers of diagnoses that are generally growing over time (except for large drops during the COVID-19 pandemic). Notably, diagnoses remained steady between 2,400-4,000 new cases per month (with some variation). Increases may be due to changes in clinical practice, e.g., the administration of more tests to ascertain whether any dementia is present, along with changes in demand for such tests due to shifts in pension rules. Both demand and supply of dementia exams may be due to awareness of the "Paisa Mutation" throughout Colombia.

Figure 2 shows the distribution of age of first diagnosis in our analysis sample (omitting controls). As would be expected, the probability of diagnosis peaks around age 80. Strikingly, a large portion of individuals are diagnosed prior to age 60. Generally, the definition of "early-onset dementia" is age 40-64 and is exceedingly rarely diagnosed. Colombia is an exception. 17.1% of diagnosis occur among individuals younger than 60; 8.9% before age 50, and 5.1% before age 40. Fully 2.4% of diagnoses occur among people between ages 18 and 30.

An obvious question is whether age patterns are driven by ADRD or non-ND. Thus, we plot diagnoses across age groups for both. For completeness, we separate ADRD into Alzheimer's Disease and Related Dimentias. The resulting age plot is presented in Figure 3. It is clear from the plot that all dementias tend to occur later. However, the lion's share of younger dementia is non-ND rather than ADRD, which aligns with our understanding of ADRD as a neurodegenerative condition largely affecting older individuals, the exception

[7]We note that further extractions of larger samples could lead to different results from those presented here; the extraction is a time-consuming process since records are often pdfs (i.e., written documents), which is why we began with a smaller sub-sample.

(also apparent in the figure) that ADRD does exist among younger groups due to the Paisa Mutation.

Figure 5 expands our main analytical sample to include first diagnoses in earlier years, including 2009-2022 and for each year replicates Figure 2. The goal is to assess whether the age distribution of first diagnoses in Figure 2 is similar across years. In general, Figure 5 shows similar patterns across years. Diagnoses peak around age 80, as we expect from other countries. However, across years, early diagnoses comprise a large proportion of total diagnoses. Moreover, we observe that the younger diagnoses seem to be rising as a proportion of total diagnoses. This could be due to increasing awareness of the Paisa mutation and thus changes to supply and demand for tests leading to possible dementia diagnoses. Changes may also reflect policy shifts allowing early retirement not only for individuals with debilitating medical conditions, but also for their caretakers, which could further incentivize people to seek medical attention for young family members, again leading to relatively more diagnoses among the very young.

Next, we consider gender differences in Figure 6. According the figure, diagnoses among younger people occur among both males and females, but youth diagnoses (at ages between 20-40) are roughly double in proportion for males compared to females. In other words, much of the pattern of diagnosed early-onset dementia is driven by males. That said, youth diagnoses of dementia are still prevalent among females; roughly 25% of total diagnoses for females occur below age 60 (which is considered early onset dementia) compared to about 40% for males. These distinct age patterns prompt us to consider heterogeneity by gender in our analyses.

Finally, we examine age of first diagnosis patterns in a specific region, Antioquia, compared to the rest of the country. The reason is that the Paisa mutation originated in Antioquia and so we would expect higher rates of youth diagnoses there. According to Figure 7. While we do see that younger diagnoses make up a larger portion of total diagnoses in Antioquia versus the rest of the country, the differences are quite. For example, it is not the case that all early (or nearly all) diagnoses among young people are concentrated in Antioquia. These patterns suggest two possibilities. First, due to migration to other states, it is possible that the Paisa mutation is more evenly distributed across Colombia than previously understood. The other possibility is that are not due to the Paisa mutation *per se*, but instead reflect general attunement to the possibility of dementia among younger individuals and thus more diagnoses, most of which would be due to non-ND.

3 Empirical Approach

The goal is to examine differences in our sample of individuals who were first diagnosed with all dementias in 2020-2022 (and who appear in both the health (RIPS) and labor (PILA) administrative data sets) compared to similar individuals who are not observed as diagnosed. The first group is the "treatment" group, and the second group is the "control" group. Here we discuss four approaches to examining differences. For a given outcome, we examine average trajectories over time, separately for treatment and control groups. Finally, we assess this measure by age groups and by gender, and separately for each outcome, to matched individuals in the control groups to understand the magnitude of losses relative to similar individuals.

3.1 Pre-diagnosis Effects

We begin by characterizing the evolution of labor force participation and log wages around the time of first diagnosis for treated and control individuals. For each group $g \in \{\text{treated}, \text{control}\}$ and event time $\tau \in \{-10, \ldots, 2\}$, defined relative to the diagnosis date, we compute group-specific mean outcomes:

$$\bar{Y}_{g\tau} = \frac{1}{N_{g\tau}} \sum_{i \in g} Y_{i\tau}, \tag{1}$$

where $N_{g\tau}$ denotes the number of individuals in group g observed at event time τ.

We present these averages graphically and report 95% confidence intervals based on the standard error of the mean for each (g, τ) cell. For control individuals, event time is assigned to coincide with the diagnosis date of their matched treated counterpart, as described in the previous section.

This descriptive analysis serves two purposes. First, it provides a transparent description of the timing and magnitude of outcome changes around diagnosis. Second, it allows an assessment of whether treated and control individuals exhibit comparable labor market trajectories prior to diagnosis. Although we do not estimate a dynamic causal model at this stage, the absence of systematic divergence in pre-diagnosis means across groups lends support to the plausibility of the identifying assumptions underlying the main empirical analysis.

3.2 Cumulative Effects

To summarize the overall magnitude of pre-diagnosis differences between treated and control groups, we calculate cumulative effects by summing observed differences over the ten-year period prior to diagnosis (event times $\tau = -10$ to $\tau = 0$). For each outcome, we compute the mean difference between treated and control individuals within each event-time period, then sum these differences across all eleven periods.

Formally, for outcome Y and subgroup s, the cumulative effect is:

$$\text{Cumulative Effect}_s = \sum_{\tau=-10}^{0} (\bar{Y}_{s,\text{treat},\tau} - \bar{Y}_{s,\text{control},\tau}) \times 12 \tag{2}$$

where $\bar{Y}_{s,g,\tau}$ represents the mean outcome for group g (treated or control) in subgroup s at event time τ. We multiply by 12 to express annual outcomes in monthly units, facilitating interpretation of the total accumulated difference over the observation window.

We compute cumulative effects for four key outcomes: healthcare utilization, log wages, log credit balance, and conviction duration (in months). We calculate these measures for the full sample and separately by gender (male, female) and age groups (40-64, 65+), allowing us to examine heterogeneity in pre-diagnosis trajectories. Statistical inference is based on t-tests comparing treated and control group means over the cumulative period, with 95% confidence intervals constructed using robust standard errors.

4 Results

Our primary analysis focuses on working-age individuals aged 40-64 at first diagnosis, as this group faces the most significant economic consequences from dementia while remaining substantially engaged in formal labor markets. We examine four key outcomes: healthcare utilization, labor supply, wages, and credit market participation.

We organize the results as follows. First, we present evidence plotting average outcomes for treated and control groups across event time relative to diagnosis. Then, we calculate cumulative pre-diagnosis effects, summarizing the total magnitude of health and economic deterioration in the decade preceding diagnosis. We present results for the full sample and separately by gender and age subgroups to examine heterogeneity in pre-diagnosis trajectories.

4.1 Health Outcomes

Figure 8 plots average log total healthcare utilization (measured in Colombian pesos) for treated and control groups aged 40-64 across event time relative to first diagnosis, separately for all dementias, ADRD, and non-ND. Healthcare utilization exhibits a distinct pattern, with treated individuals showing substantially higher utilization throughout the entire pre-diagnosis period.

Ten years before diagnosis, working-age treated individuals already demonstrate higher healthcare utilization than controls. For all dementias combined, treated individuals average approximately 1.8 log points in healthcare spending at $t = -10$, compared to 1.6 log points for controls—a difference of roughly 22%.[8] This gap persists and grows steadily over the subsequent decade.

The divergence accelerates sharply in the final years before diagnosis. At $t = -1$, treated individuals' healthcare utilization peaks at approximately 5.0 log points, compared to 3.0 log points for controls, representing more than a ten-fold difference in spending levels. Post-diagnosis, treated individuals' utilization reaches 7.0 log points by $t = 1$ and remains substantially elevated relative to controls, whose utilization stays relatively stable around 3.2-3.5 log points throughout the observation window. The patterns are consistent across ADRD and non-ND diagnoses.

These patterns reveal two key insights. First, the higher healthcare utilization present even a decade before diagnosis suggests that individuals who will eventually receive a dementia diagnosis experience health problems requiring medical attention years before cognitive symptoms become severe enough for formal diagnosis. This may reflect early prodromal symptoms, comorbid conditions, or other health issues that presage cognitive decline. Second, the dramatic spike in utilization immediately before diagnosis likely reflects the diagnostic process itself, involving specialist consultations, cognitive testing, and imaging studies required to confirm dementia. The post-diagnosis decline suggests a shift from diagnostic to maintenance care, though utilization remains substantially elevated relative to controls, consistent with the ongoing medical needs of dementia patients.

Appendix A.1 presents the evolution of total medical visits over the full event window ($t = -10$ to $t = 2$), with gender-specific estimates shown in Appendix A.2 (females) and Appendix A.3 (males).

[8] The percentage difference is calculated as $e^{1.8-1.6} - 1 \approx 0.22$.

4.1.1 Network Analysis of Comorbidities

We employ network analysis to characterize the structure of comorbidities in the year preceding ADRD diagnosis, following methodological approaches established in the medical networks literature (Barabási et al., 2011; Jensen et al., 2014; Dervić et al., 2025). This framework allows us to identify which diagnostic conditions are central to disease progression and to quantify the strength of comorbidity relationships across the patient population.

Our network construction proceeds in two stages. First, we obtained the raw patient-diagnosis relationships as a bipartite network, where one set of nodes represents patients and another represents diagnostic categories. An edge connects a patient to a diagnosis if that patient received the diagnosis during the observation period. This bipartite structure captures the fundamental many-to-many relationship between patients and diseases: each patient may have multiple diagnoses, and each diagnosis appears across multiple patients.

We then project this bipartite network onto the diagnosis space to construct a diagnosis co-occurrence network. In this projection, nodes represent ICD-10 diagnostic categories, and an edge connects two categories if they co-occur in the same patient. The edge weight w_{ij} between diagnoses i and j equals the number of unique patients who received both diagnoses during the observation period. This projection transforms patient-level information into a disease-level representation that reveals the structure of comorbidity patterns (Hidalgo et al., 2009).

Formally, let $P = \{p_1, \ldots, p_N\}$ denote the set of patients and $D = \{d_1, \ldots, d_M\}$ the set of diagnostic categories. The bipartite network is represented by an $N \times M$ matrix $\mathbf{B}$, where $B_{pk} = 1$ if patient p received diagnosis d_k, and $B_{pk} = 0$ otherwise. The diagnosis co-occurrence network is then defined by the adjacency matrix $\mathbf{A} = \mathbf{B}^T\mathbf{B}$, where A_{ij} counts the number of patients with both diagnoses i and j.

To aggregate diagnoses at the ICD-10 chapter level (first-level classification), we reduce the diagnostic space from thousands of individual codes to 19 clinically meaningful categories. This aggregation serves multiple purposes: it reduces noise from rare diagnosis combinations, focuses analysis on clinically interpretable categories, and facilitates visual interpretation of network structure. We exclude three ICD-10 chapters from the analysis: Chapter XVIII (Symptoms, Signs, and Abnormal Clinical Findings), Chapter XX (External Causes of Morbidity and Mortality), and Chapter XXI (Factors Influencing Health Status), as these represent administrative categories or non-specific presentations rather than established diagnoses.

Then the comorbidity structure is characterized using three complementary measures: prevalence, intensity, and network centrality.

Prevalence (Incidence) For each diagnostic category c, we define prevalence as the percentage of patients receiving at least one diagnosis in that category:

$$\text{Prevalence}_c = \left(\frac{N_c^{\text{unique}}}{N^{\text{total}}} \right) \times 100 \tag{3}$$

where N_c^{unique} is the number of unique patients with at least one diagnosis in category c, and N^{total} is the total number of patients in the sample. This measure treats each patient equally regardless of diagnosis frequency and captures the population-level burden of each diagnostic category.

Diagnostic Intensity We measure diagnostic intensity as the average number of diagnoses per affected patient:

$$\text{Intensity}_c = \frac{D_c^{\text{total}}}{N_c^{\text{unique}}} \tag{4}$$

where D_c^{total} is the total number of diagnosis instances in category c (including repeated diagnoses). Higher intensity indicates either greater disease severity, poorer disease control, or more chronic conditions requiring frequent medical attention. For instance, an intensity of 2.5 indicates that affected patients received, on average, 2.5 diagnoses in this category during the observation period.

Centrality Measures We employ eigenvector network centrality to quantify the structural importance of each diagnostic category within the comorbidity network. *Eigenvector centrality* assesses the importance of a diagnostic category based on the importance of its neighbors:

$$\text{Eigenvector}_c = \frac{1}{\lambda} \sum_{j \in M(c)} A_{cj} \cdot \text{Eigenvector}_j \tag{5}$$

where λ is the largest eigenvalue of the adjacency matrix $\mathbf{A}$, $M(c)$ is the set of neighbors of category c, and A_{cj} is the edge weight between categories c and j (number of shared patients). Eigenvector centrality assigns higher scores to categories connected to other highly connected categories. A diagnostic category can have high eigenvector centrality even with moderate prevalence if it co-occurs with highly prevalent conditions, identifying diagnoses embedded within dense clusters of comorbidities. For example, diabetes often displays high eigenvector centrality due to its connections to hypertension and cardiovascular diseases, which themselves are highly connected.

Finally, we visualize diagnosis co-occurrence networks in Figure A.4 using force-directed layout algorithms that position nodes based on their connectivity patterns. Node size is proportional to prevalence (percentage of patients with the diagnosis), and edge width indicates the number of patients sharing both diagnoses. Node colors represent ICD-10 chapters. Node positions are determined using a combination of closeness and eigenvector centrality measures to emphasize structurally important diagnoses. This visualization approach, following Jensen et al. (2014), reveals both the prevalence of individual conditions and the structure of their interconnections.

Table 4 presents the comorbidity structure for treated and control groups in the year before diagnosis, measured by prevalence (percentage of patients with at least one diagnosis in each category), diagnostic intensity (average diagnoses per affected patient), and eigenvector centrality (which identifies categories embedded within dense comorbidity clusters). The most striking difference appears in mental health diagnoses, which affect 41.6% of the treated group compared to only 10.4% of controls—a 31.2 percentage point gap. This pattern likely reflects both the prodromal psychiatric symptoms common in early dementia and the diagnostic challenges in distinguishing cognitive decline from primary psychiatric disorders. Nervous system diagnoses show a similar pattern, affecting 29.5% of treated patients versus 11.6% of controls, consistent with the neurological basis of dementia and potentially reflecting earlier diagnoses of conditions such as Parkinson's disease or cerebrovascular disease that share pathophysiological mechanisms with dementia.

Beyond prevalence differences, the network centrality measures reveal how these conditions are embedded within broader comorbidity patterns. Mental health diagnoses exhibit substantially higher eigenvector centrality in the treated group (0.883 versus 0.268), indicating they co-occur with other highly connected diagnostic categories. This suggests mental health conditions in the pre-diagnosis period are part of a dense cluster of interrelated health problems rather than isolated diagnoses.

Given the prominence of mental health diagnoses in distinguishing treated from control groups, Appendix Table A.1 and Figure A.4 present a detailed analysis restricted to ICD-10 Chapter F (Mental and Behavioural Disorders). The appendix network visualization and accompanying table decompose the mental health category into its constituent diagnoses, revealing that addiction-related disorders account for the majority of the treatment-control gap in mental health prevalence.

4.2 Labor Market Outcomes

Figures 10 and 11 plot average log wages and formal labor supply for treated and control groups in this age range across event time relative to first diagnosis, separately for all dementias, ADRD, and non-ND. The horizontal axis represents years until diagnosis, where -10 indicates ten years prior to diagnosis and 0 marks the diagnosis year. Results for the full sample (including those 65+) are presented for comparison.

Both outcomes exhibit strikingly similar patterns across dementia types. Ten years before diagnosis, treated and control groups aged 40-64 are nearly indistinguishable in both earnings and employment. However, divergence begins approximately six to seven years prior to diagnosis, with gaps expanding steadily as diagnosis approaches.

Log wages follow an analogous trajectory. At $t = -10$, both groups earn similar wages across all age categories. By one year before diagnosis, the control group averages a log wage of approximately 3.90 compared to 3.60 for the treated group, representing a 35% wage difference.[9] This gap persists and widens post-diagnosis. The patterns are consistent across both ADRD and non-ND diagnoses, though the 40-64 age group shows larger absolute gaps due to higher baseline employment and earnings.

For labor supply, employment rates among working-age individuals (40-64) peak around 55-56% at $t = -5$ for all groups. By one year before diagnosis, the control group maintains approximately 54% employment while the treated group declines to approximately 52%, a 2 percentage point difference. The decline accelerates post-diagnosis, with treated individuals' employment falling to roughly 40% by two years after diagnosis. Among older individuals (65+), employment rates are lower overall but exhibit similar divergence patterns, declining from approximately 45% at baseline to 37% at diagnosis for treated individuals versus 40% for controls.

These patterns reveal substantial labor market deterioration in the years preceding diagnosis. Critically, the near-zero gaps at $t = -10$ indicate that selection on permanent productivity differences cannot fully explain the observed divergence. If individuals eventually diagnosed were inherently less productive, we would observe persistent gaps throughout the pre-diagnosis period. Instead, the gradual emergence and growth of gaps beginning seven to eight years before diagnosis indicates that progressive health deterioration—rather than fixed individual differences—drives the labor market consequences we observe. This interpretation aligns with clinical evidence that cognitive decline precedes formal dementia diagnosis by several years, during which individuals experience increasing difficulty maintaining employment and productivity.

[9] The percentage difference is calculated as $e^{3.90-3.60} - 1 \approx 0.35$.

The same analysis disaggregated by gender is presented in Appendix for wages in A.5, A.6 and labor supply A.7, A.8 .

4.3 Credit Outcomes

Figure 12 plots average log total credit balance for treated and control groups. In contrast to healthcare utilization and labor market outcomes, credit balances show minimal divergence between treated and control groups throughout the pre-diagnosis period.

Ten years before diagnosis, treated and control individuals aged 40-64 hold similar credit balances across all dementia types. For all dementias combined, both groups average approximately 12.3 log points at $t = -10$, effectively indistinguishable. Both groups exhibit gradual growth in credit balances over time, reaching approximately 13.0-13.2 log points by five years before diagnosis, with treated and control groups tracking each other closely throughout this period.

In the final years before diagnosis, credit balances continue to rise for both groups, though a modest gap begins to emerge and persists post-diagnosis, with both groups maintaining relatively stable credit levels around 13.5-13.8 log points through $t = 2$.

The patterns are remarkably consistent across ADRD and non-ND diagnoses, with both showing nearly identical trajectories and minimal differences between treated and control groups. The slight widening of gaps in the immediate pre-diagnosis period may reflect reduced borrowing capacity as labor market attachment weakens, though the effects are economically small compared to the substantial deterioration observed in employment and wages.

The main credit estimates, disaggregated by gender, are presented in Appendix A.9 and Appendix A.10. Appendix A.11 reports the total number of credits at each event time from $t = -10$ to $t = 2$.

4.4 Crime Outcomes

Figure 13 depicts the evolution of average conviction rates, defined as the proportion of individuals with an active criminal conviction, for treated and control individuals aged 40–64, indexed relative to the year of first diagnosis ($t = 0$). The figure reports results separately for all dementias, ADRD, and non-ND.

Conviction rates are low throughout the observation window for both groups. Ten years prior to diagnosis ($t = -10$), rates average approximately 0.008 (0.8 percent) among treated individuals and 0.006 (0.6 percent) among controls across all dementia types, indicating

a small but persistent baseline difference. Over the subsequent pre-diagnosis period, conviction rates remain stable, with treated individuals consistently exhibiting slightly higher rates—typically around 0.008 to 0.010, while control rates remain between 0.006 and 0.007 through $t = -5$.

In the years immediately preceding diagnosis, conviction rates continue to evolve smoothly for both groups, with no evidence of sharp divergence. At $t = -1$, treated individuals exhibit conviction rates of approximately 0.010 (1.0 percent), compared to 0.007 (0.7 percent) for controls. Post-diagnosis patterns are similarly stable: treated individuals maintain conviction rates around 0.010–0.011, while control rates remain between 0.007 and 0.008 through $t = 2$.

These dynamics are remarkably consistent across ADRD and non-ND diagnoses, with nearly identical trajectories and no discernible discontinuities at the time of diagnosis. The persistence of a small gap throughout the entire observation period suggests that higher conviction rates among treated individuals reflect pre-existing differences rather than changes associated with the onset or progression of cognitive impairment.

Taken together, these findings indicate that criminal justice involvement, as measured by active convictions, does not exhibit meaningful changes as individuals approach or pass the point of dementia diagnosis. This contrasts sharply with labor market outcomes, where treated and control individuals diverge substantially in the years leading up to diagnosis, and suggests that dementia-related cognitive decline does not manifest through increased criminal behavior in this working-age population.

The same analysis disaggregated by gender is presented in Appendix A.12 and A.13.

4.5 Summarizing Aggregate Effects

To get a sense of the total effects of pre-diagnosis dementia, we conduct the following exercise. Essentially, we sum the differences between the treatment and control groups for different groups, starting at the point when the gap between the two becomes significant. We do this for people of all ages and then separately for people ages 40-64 and again for people aged 65 or more. We also perform the exercise separately for males and females. This allows us to provide a single figure of all results and to compare magnitudes of total differences between diagnosed and never diagnosed across groups and outcomes. Results are presented in Figure 14. For example, the first panel shows that healthcare utilization rose for all ages and genders. However, the largest increase is among those aged 40-64 versus people 65 and older. The second panel shows that wages declined for all groups except for women aged 40-64. The third panel shows that impacts on credit are nearly negligible except for

women age 65 and over. Finally, we see that crime shows mixed effects that are driven, unsurprisingly, by individuals aged 40-64. Somewhat surprisingly, and worth attention in future analyses, we see that men exhibit reductions in criminal activity, while women exhibit increases. Again, the idea is that the cumulative difference criminal activity for women with a diagnoses, compared to women without one, is an increase. For men, it is a decrease. Likely these differences reflect differences in selection into criminality across genders, which would lead to different relationships with dementia.

The same analysis disaggregated by sub-diagnosis is presented in Appendix A.14 and A.15.

5 Conclusion

Research on dementia from the past several years has shown that ADRD is likely to be a larger problem than previously understood. It affects people on more dimensions and starting at an earlier age that earlier literature had reported. Moreover, the consequences for dementia extend to people who are never diagnosed. A full accounting of the impacts of dementia, however, is not limited to ADRD. Rather, non-ND is another problem that is far more likely to affect younger people.

Together, the idea that ADRD affects people earlier than understood along with consideration of non-ND suggest that a focus solely on the elderly leads to an incomplete accounting of dementia and its consequences. For example, if we focus on outcomes associated with old age, such as retirement wealth, we can miss some of the costs associated with dementia.

This paper examines all forms of dementia diagnoses among Colombians across the age distribution, with particular focus on working-age Colombians. The goal is to develop a more complete understanding of the full range of population implications of different forms of dementia. Our results are preliminary and much work remains to be done. However, we are able to show several intriguing patterns. First, in line with other literature focussed on ADRD, we also find that the consequences of dementia emerge long before diagnosis, both for ADRD and for non-ND and for both the elderly and relatively young people. Second, addiction is an important co-morbidity for dementia, again across age groups and for both ADRD and non-ND. Third, extending the outcomes we consider, we find novel evidence that dementia and criminality are linked, which opens up new questions about the causes and consequences of limited cognitive function. In future analyses, we will explore these relationships to better understand the mechanisms that underlie them.

References

Acosta-Baena, Natalia, Diego Sepulveda-Falla, Carlos Mario Lopera-Gómez, Mario César Jaramillo-Elorza, Sonia Moreno, Daniel Camilo Aguirre-Acevedo, Amanda Saldarriaga, and Francisco Lopera, "Pre-dementia clinical stages in presenilin 1 E280A familial early-onset Alzheimer's disease: a retrospective cohort study," *The Lancet Neurology*, 2011, *10* (3), 213–220.

Aguirre-Acevedo, Daniel C, Francisco Lopera, Diana Henao, Valentina Tirado, Joseph F Arboleda-Velasquez, María C Franco, César Tobón, Laura Piedrahita, Sergio Moreno, Claudia Muñoz et al., "Cognitive Decline in a Colombian Kindred With Autosomal Dominant Alzheimer Disease: A Retrospective Cohort Study," *JAMA Neurology*, 2016, *73* (3), 431–438.

Albanesi, Stefania and Jiyeon Kim, "Effects of the COVID-19 Recession on the US Labor Market: Occupation, Family, and Gender," *Journal of Economic Perspectives*, 2021, *35* (3), 3–24.

Barabási, Albert-László, Natali Gulbahce, and Joseph Loscalzo, "Network medicine: a network-based approach to human disease," *Nature reviews genetics*, 2011, *12* (1), 56–68.

Bernal, Raquel, Sandra Caro, and Ximena Peña, "The effects of a long maternity leave on maternal employment: Evidence from a natural experiment in Colombia," *CEDE Working Paper*, 2017.

Blue Cross Blue Shield Association, "Early-Onset Dementia and Alzheimer's Diagnoses Spiked 373 Percent for Generation X and Millennials," *The Health of America Report Series*, 2020.

Borgbjerg, Anne Katrine, Esben Agerbo, Nabanita Datta Gupta, and Timothy J Halliday, "Genetic Predictors of Cognitive Decline and Labor Market Exit," Technical Report DP17714, IZA Institute of Labor Economics 2024.

Brown, T. T., J. A. Pagan, and E. Bastida, "The impact of diabetes on employment and work productivity," *Diabetes Care*, 2005, *28* (11), 2662–2667.

Buitrago, Francisco, Ana García, and Rodrigo Zárate, "Uso de datos administrativos para evaluar políticas públicas en salud en Colombia," *Revista de Economía Institucional*, 2021, *23* (45), 105–130.

Camacho, Adriana and Emily Conover, "Effects of subsidies on health: Evidence from Colombia's health insurance system," *World Bank Economic Review*, 2013, *27* (3), 426–451.

Chandra, Amitabh, Courtney Coile, and Corina Mommaerts, "What can economics say about Alzheimer's disease?," *Journal of Economic Literature*, 2023, *61* (2), 428–470.

Cronin, Christopher, Matthew Forsstrom, and Nicholas W. Papageorge, "What Good are Treatment Effects Without Treatment? Mental Health and the Reluctance to Use Talk Therapy," *NBER Working Paper*, 2020, (w27711).

DANE, "Empleo informal y seguridad social," 2024. Accessed: 2025-07-31.

Dervić, Elma, Katharina Ledebur, Stefan Thurner, and Peter Klimek, "Comorbidity networks from population-wide health data: aggregated data of 8.9 M hospital patients (1997–2014)," *Scientific Data*, 2025, *12* (1), 215.

French, Michael T, M Caroline Roebuck, and Paul K Alexandre, "Drug use, drug abuse, and labour market outcomes," *The Economic Journal*, 1997, *107* (441), 58–69.

Harris, Kathleen Mullan and Carolyn Tucker Halpern, "The National Longitudinal Study of Adolescent to Adult Health (Add Health): An Underused Resource for Developmental Science," *Annual Review of Developmental Psychology*, 2022, *4*, 297–320.

Hidalgo, César A, Nicholas Blumm, Albert-László Barabási, and Nicholas A Christakis, "A dynamic network approach for the study of human phenotypes," *PLoS computational biology*, 2009, *5* (4), e1000353.

Jensen, Anders Boeck, Pope Moseley, Tudor Oprea, Sabrina Ellesøe, Robert Eriksson, Henriette Schmock, Peter Jensen, Lars Jensen, and Søren Brunak, "Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients," *Nature communications*, 06 2014, *5*, 4022.

Jeong, Yeongmi, Nicholas W. Papageorge, Meghan M. Skira, and Kevin Thom, "Genetic Risk for Alzheimer's Disease and Related Dementias: Cognition, Economic Behavior, and Long-Run Planning," NBER Working Paper 32181 July 2025.

Lang, Linda, Angela Clifford, Li Wei, Dongmei Zhang, Daryl Leung, Glenda Augustine, Isaac M. Danat, Weiju Zhou, John R. Copeland, Kaarin J. Anstey, and Ruoling Chen, "Prevalence and determinants of undetected dementia in the community: a systematic literature review and a meta-analysis," *BMJ Open*, 2017, *7* (2).

Li, Jing, Amy Kelley, Kathleen McGarry, Lauren Hersch Nicholas, and Jonathan Skinner, "Dementia and Long-Run Trajectories in Household Finances," *National Bureau of Economic Research Working Paper*, 2022, (29772).

_ , Amy S Kelley, Kathleen McGarry, Lauren Hersch Nicholas, Nicholas W Papageorge, Jonathan Skinner, Kevin Thom, and Meghan Skira, "Genetic Risk for Alzheimer's Disease and Related Dementias: Cognition, Economic Behavior, and Actionable Information," Working Paper w32181, National Bureau of Economic Research 2024.

_ , Jonathan S. Skinner, Kathleen McGarry, Lauren Hersch Nicholas, Shao-Pang Wang, Evan Bollens-Lund, and Amy S. Kelley, "Declines in wealth among US older adults at risk of dementia," *JAMA Neurology*, 2023, *80* (11), 1250–1252.

Mazzonna, Fabrizio and Franco Peracchi, "Are Older People Aware of Their Cognitive Decline? Misperception and Financial Decision-Making," *Journal of Political Economy*, 2024, *132* (6), 1793–1830.

MINSALUD, "Guía para el reconocimiento de pensiones por invalidez," 2020. Disponible en: `https://www.minsalud.gov.co/`.

Nicholas, Lauren Hersch, Kenneth M. Langa, Julie P. W. Bynum, and Joanne W. Hsu, "Financial presentation of Alzheimer disease and related dementias," *JAMA Internal Medicine*, 2021, *181* (2), 220–227.

Papageorge, Nicholas W., "Modeling Behavior During a Pandemic: Using HIV as an Historical Analogy," *NBER Working Paper*, 2021, (w28898).

Polsky, Daniel and Anirban Basu, "Selection Bias in Observational Data," in "The Elgar Companion to Health Economics, Second Edition," Edward Elgar Publishing, 2012, chapter 46.

Shen, Karen, "Who benefits from public financing of home-based long term care? Evidence from Medicaid," *Journal of Public Economics*, 2024, *236*, 105151.

Thirumurthy, Harsha, Joshua Graff Zivin, and Markus P Goldstein, "The Economic Impact of AIDS Treatment: Labor Supply in Western Kenya," *The American Economic Review*, 2008, *98* (2), 226–231.

Zimmerman, Frederick J. and Wayne Katon, "Socioeconomic status, depression disparities, and financial strain: what lies behind the income-depression relationship?," *Health Economics*, 2005, *14* (12), 1197–1215.

Tables and Figures

Table 1: Samples

	RIPS	RIPS + PILA	ID
Panel A: Overall Sample (2009-2022)			
Total Observed	49,495,515	26,840,329	15,296,051
ADRD	592,996	164,110	126,476
Non Neurodegenerative Dementia	301,727	89,433	69,254
All Dementias	**894,723**	**253,534**	**195,730**
Panel B: First Diagnosis (2020-2022)			
ADRD	65,380	17,911	13,462
Non Neurodegenerative Dementia	78,587	26,959	19,677
All Dementias	**143,967**	**44,870**	**33,139**
Panel C: First Diagnosis (2020-2022) & Specialists & Exams & WO Urgencies			
ADRD	32,197	9,129	5,410
Non Neurodegenerative Dementia	38,064	13,436	8,326
All Dementias	**70,261**	**22,565**	**13,736**

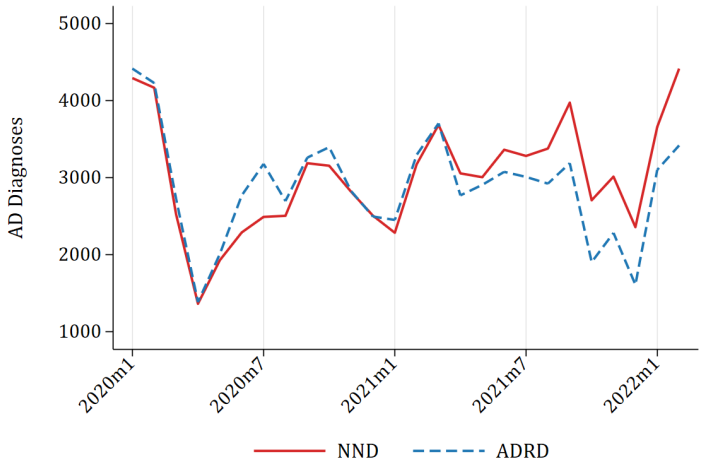
Notes: This table reports summarizes the number of observations available in our administrative data sources. RIPS corresponds to the national health services registry (*Registro Individual de Prestación de Servicios de Salud*), and PILA corresponds to formal labor market records (*Planilla Integrada de Liquidación de Aportes*). Panel A reports the total number of observations for each data source, including the number of ADRD and non-ND diagnoses. Panel B restricts the sample to individuals with a first diagnosis of Alzheimer's disease or dementia between 2020 and 2022. Panel C reports patients with relevant specialists visits and exams in the year before diagnoses, as well as diagnoses from sources different to ER visits. Diagnoses are identified using ICD-10 codes. The study sample includes individuals for whom health and labor data can be linked using unique identifiers.

Table 3: Balance Tests: Treated and Control Groups

	Pre-Diag				Post-Diag			
	Treated		Control		Treated		Control	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Panel A: Healthcare Utilization (RIPS)								
Age	74.552	11.974	74.552	11.974	74.497	12.013	74.497	12.013
Female	0.384	0.486	0.384	0.486	0.383	0.486	0.383	0.486
Total Util.	94,409,662	714,925,989	67,057,181	580,659,925	249,986,175	105,4674,019	106,135,498	645,995,584
Log of Total Util.	3.289	5.267	2.753	4.930	6.112	6.039	3.900	5.511
Total Health Events	1.240	4.916	0.909	3.755	3.592	11.597	1.623	6.549
Observations	9,247,200		18,494,400		850,738		1,701,476	
Panel B: Labor Market Outcomes (PILA)								
Age	68.436	12.771	68.103	12.768	68.335	12.811	67.991	12.815
Female	0.454	0.498	0.453	0.498	0.449	0.497	0.448	0.497
Labor Supply	0.461	0.499	0.480	0.500	0.374	0.484	0.418	0.493
Salary	552,038,687	3,117,544,759	651,694,451	4,077,900,676	623,225,302	2,387,804,534	934,088,966	4,231,242,357
Log Salary	6.218	6.817	6.498	6.864	5.177	6.770	5.853	6.987
Observations	2,929,429		5,581,202		269,439		513,016	
Panel C: Credit Market Outcomes								
Age	66.724	14.561	66.726	14.557	66.768	14.615	66.768	14.610
Female	0.484	0.500	0.484	0.500	0.483	0.500	0.483	0.500
Total Number of Credits	0.360	0.480	0.360	0.480	0.337	0.473	0.360	0.480
Total Balance	49,967,258,658	376,700,000	48,745,320,533	366,800,000	59,092,953.744	292,100,000	66,158,804,137	4426,500,000
Log of Total Balance	7.874	8.192	7.866	8.191	7.688	8.348	7.980	8.384
Observations	180,554		361,400		30,510		61,064	
Panel D: Criminal Justice Outcomes								
Age	65.189	12.865	65.189	12.865	65.117	12.847	65.117	12.847
Female	0.487	0.500	0.487	0.500	0.486	0.500	0.486	0.500
Time convicted	0.004	0.062	0.004	0.061	0.007	0.081	0.006	0.078
Ever convicted	0.016	0.124	0.015	0.121	0.016	0.124	0.014	0.119
Exact date of conviction	0.000	0.007	0.000	0.006	0.000	0.002	0.000	0.003
Observations	1,252,933		2,505,866		167,622		335,244	

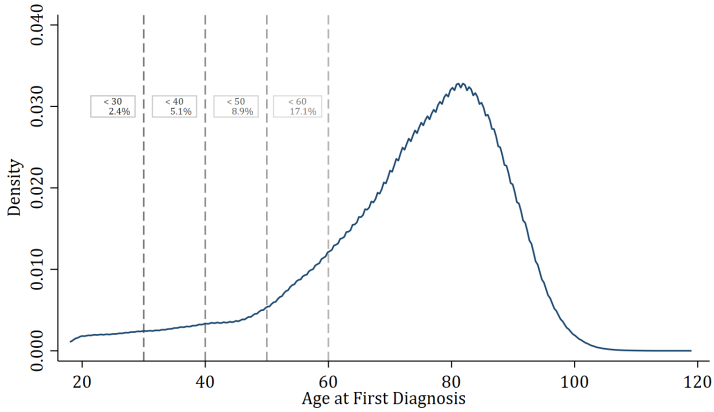
Notes: This table presents summary statistics for treated individuals (diagnosed with ADRD or non-ND) and matched control individuals across four outcome domains. Pre-Event refers to the period before ADRD or non-ND diagnosis, while Post-Event refers to the period after diagnosis. Panel A shows healthcare utilization measures from RIPS administrative records. Panel B presents labor market outcomes from PILA employment records. Panel C displays credit market variables from financial administrative data. Panel D reports criminal justice outcomes. Controls are matched to treated individuals on age, sex, municipality, and time period.

Figure 1: New Diagnosis by Month: Alzheimer's and Dementia



Notes: The figure plots the monthly number of first diagnoses for dementia and Alzheimer's disease from 2020 to 2022 using national administrative health records. Each point reflects the number of new cases recorded in a given calendar month. The sample includes all individuals with a first-time diagnosis during the period. Dementia diagnoses are plotted on the left y-axis and Alzheimer's diagnoses on the right y-axis.

Figure 2: Age Distribution at First Diagnosis



Notes: The figure displays the distribution of age at first diagnosis for individuals with dementia or Alzheimer's disease between 2020 and 2022. The density is estimated using kernel smoothing. Vertical lines indicate cumulative shares of the sample diagnosed before specific age thresholds: 7.0% under age 30, 12.8% under 40, 19.1% under 50, and 29.4% under 60.

Figure 3: Ages Distributions by Patient Group

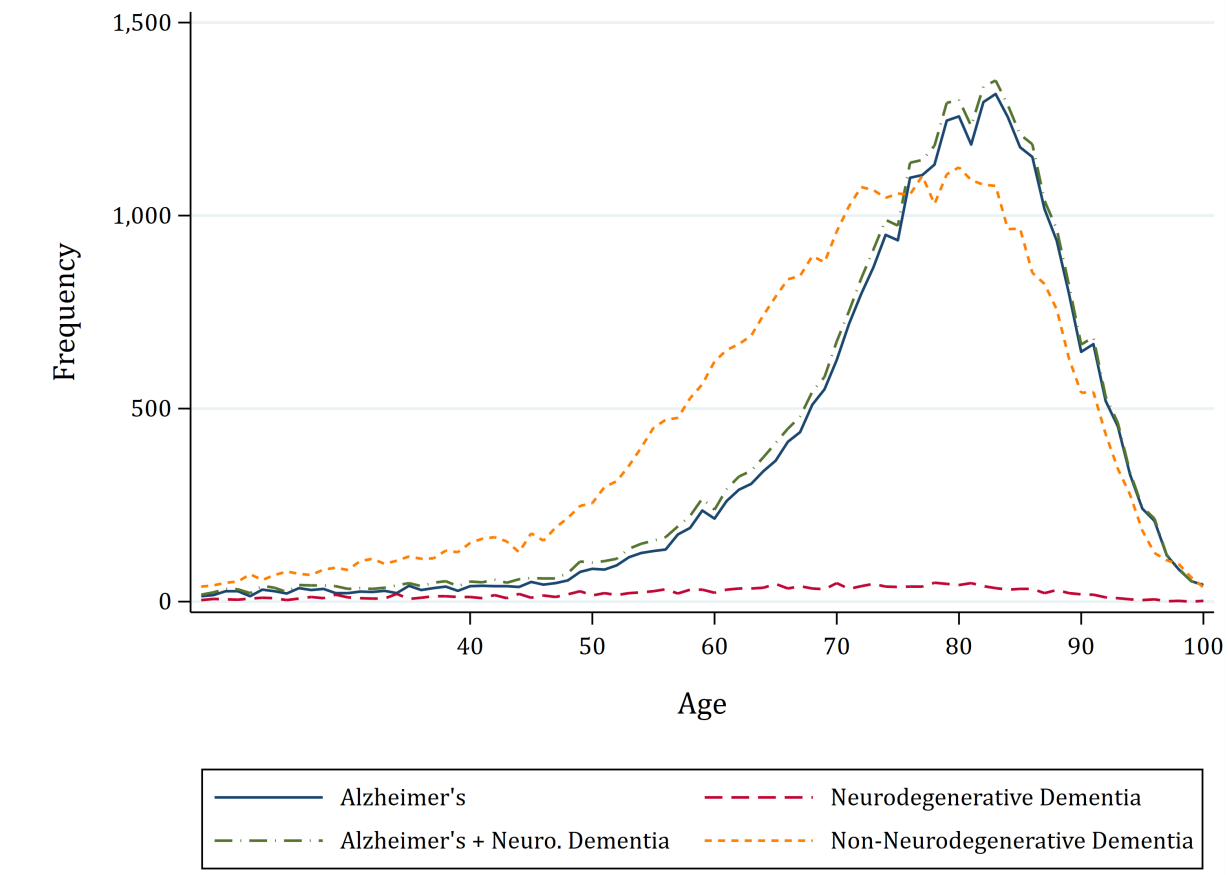
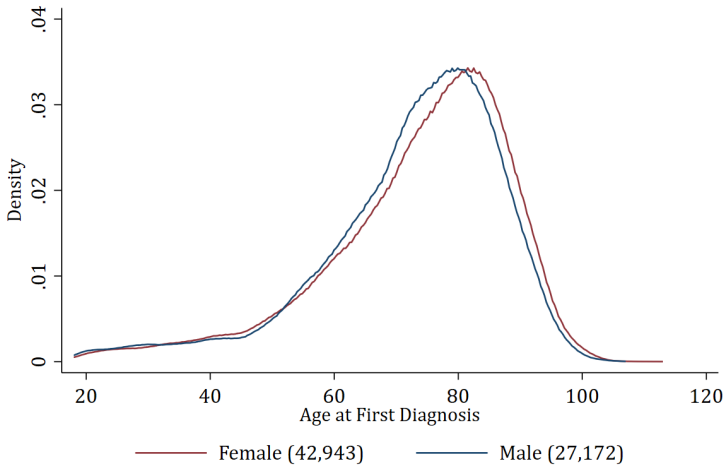
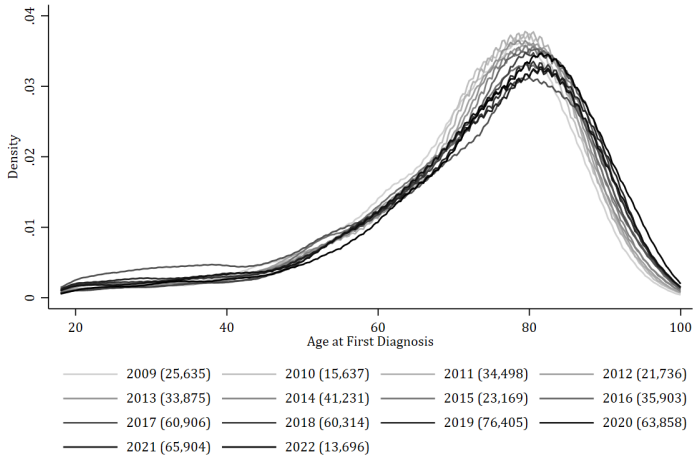


Figure 4: Age Distribution at First Diagnosis: Males vs Females



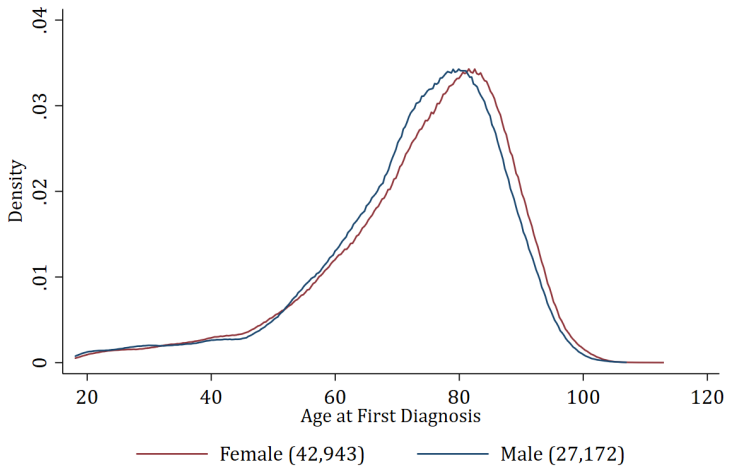
Notes: The figure displays kernel density estimates of age at first diagnosis for individuals diagnosed with dementia or Alzheimer's disease between 2020 and 2022, comparing males to females. Densities are estimated using a common smoothing bandwidth. Sample sizes are shown in the legend. The relatively younger age distribution observed in males may be related to the presence of early-onset Alzheimer's cases in that sex, including the well-documented GNA family cluster, although we do not observe genetic data directly.

Figure 5: Age Distribution at First Diagnosis by year



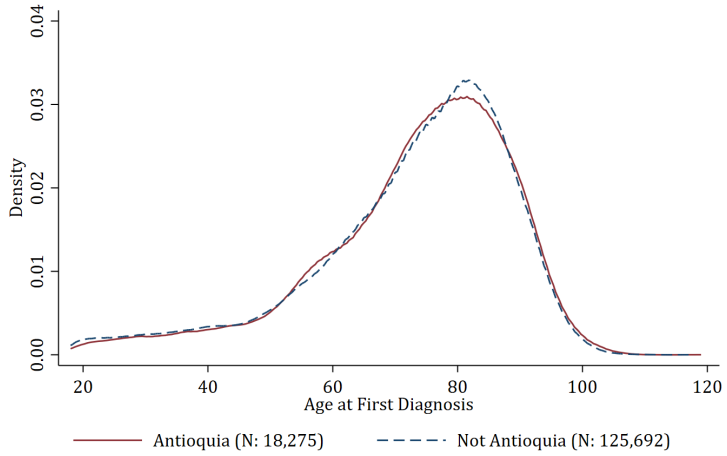
Notes: This figure shows the distribution of age at first diagnosis for individuals with ADRD or non ND diagnoses, separately by year from 2009 to 2022. Each line corresponds to a kernel density estimate for a given calendar year, using a common smoothing bandwidth. The total number of first diagnoses per year is shown in the legend.

Figure 6: Age Distribution at First Diagnosis: Males vs Females



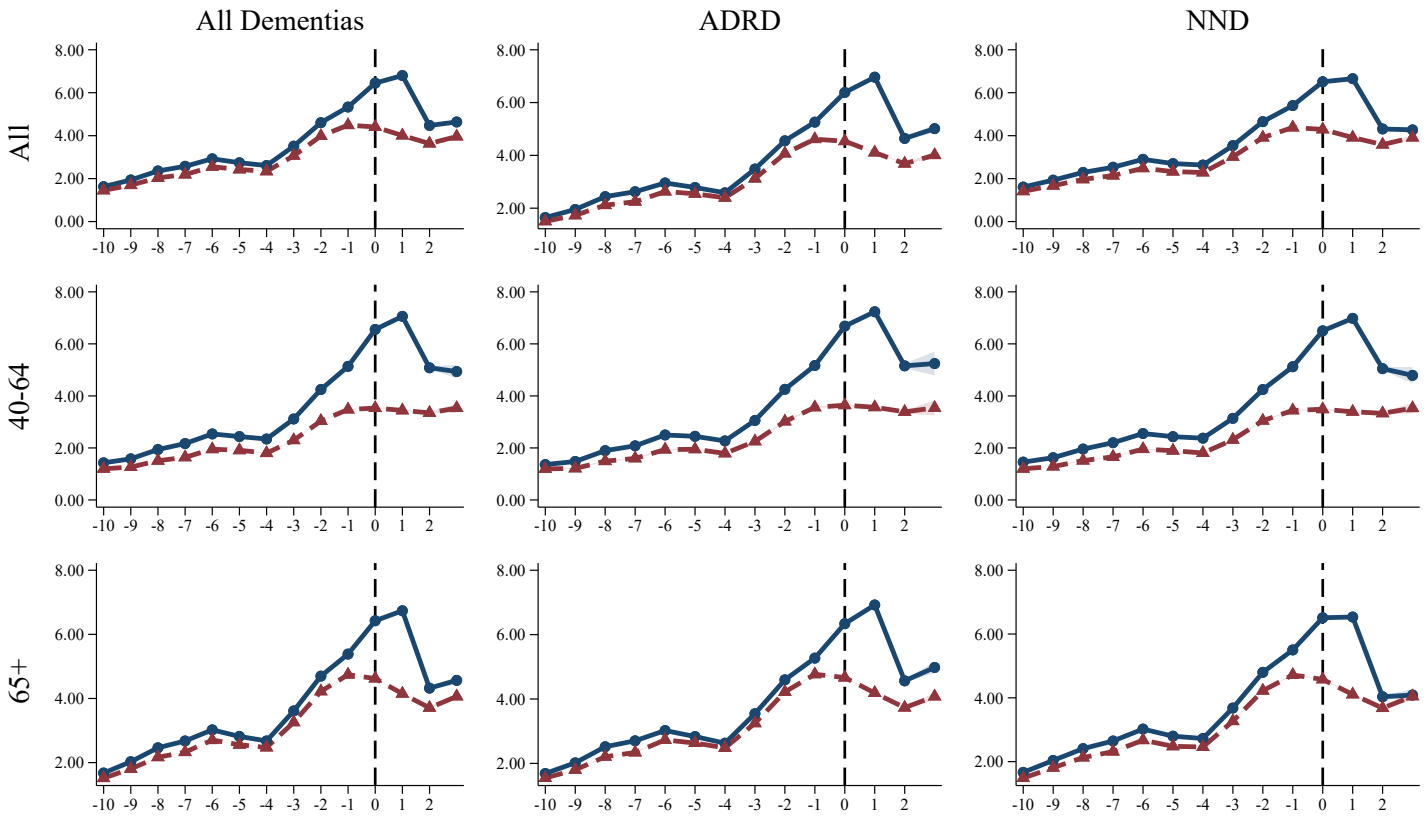
Notes: The figure displays kernel density estimates of age at first diagnosis for individuals diagnosed with all dementias between 2020 and 2022, comparing males to females. Densities are estimated using a common smoothing bandwidth. Sample sizes are shown in the legend. The relatively younger age distribution observed in males may be related to the presence of early-onset Alzheimer's cases in that sex, including the well-documented GNA family cluster, although we do not observe genetic data directly.

Figure 7: Age Distribution at First Diagnosis: Antioquia vs Rest of Colombia



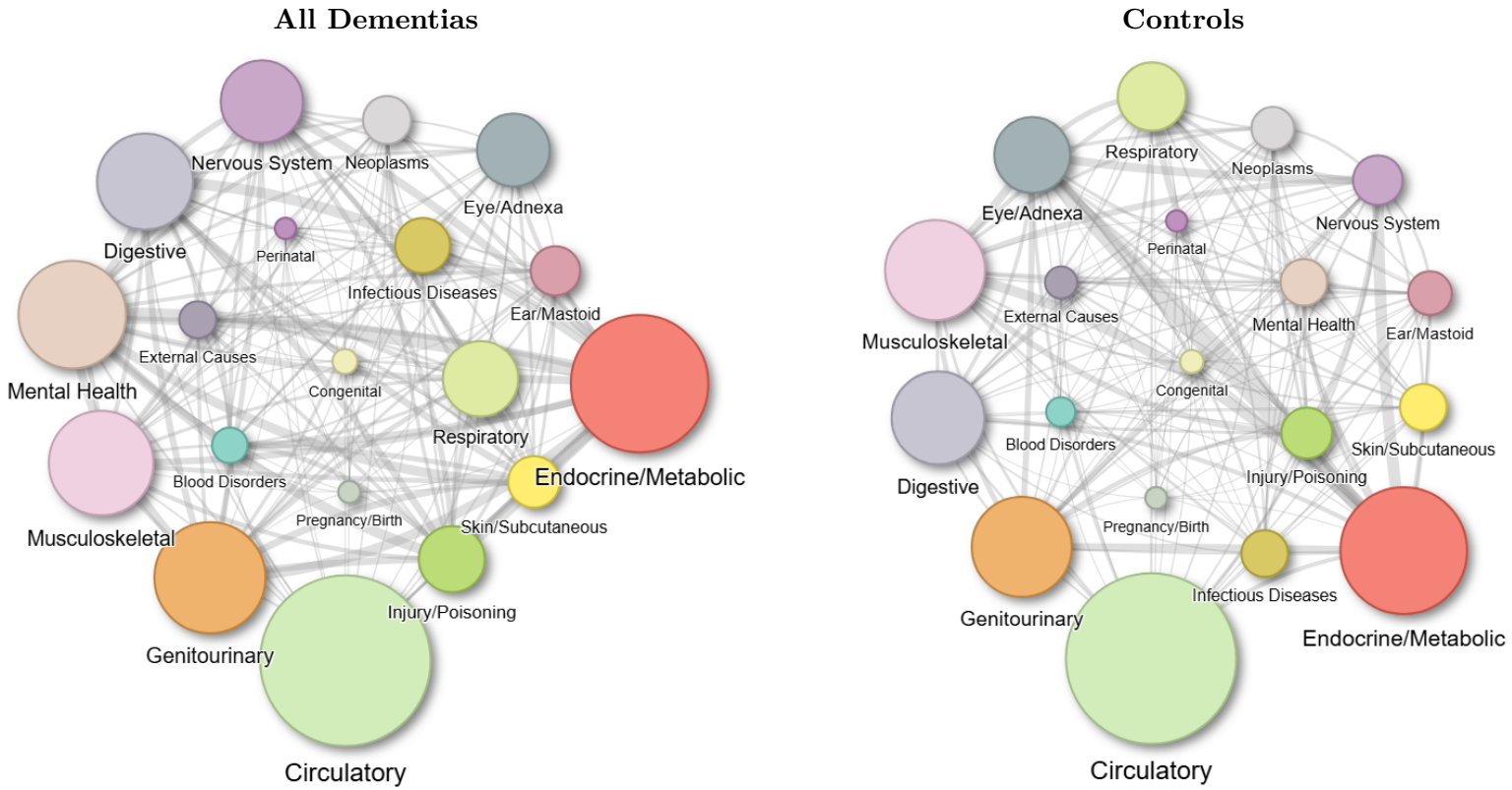
Notes: The figure displays kernel density estimates of age at first diagnosis for individuals diagnosed with all dementias between 2020 and 2022, comparing Antioquia to the rest of the country. Densities are estimated using a common smoothing bandwidth. Sample sizes are shown in the legend. The relatively younger age distribution observed in Antioquia may be related to the presence of early-onset Alzheimer's cases in the region, including the well-documented GNA family cluster, although we do not observe genetic data directly.

Figure 8: Health Outcomes of patients and controls: Log of Total Utilization



Notes: The figure displays log total utilization (in Colombian pesos) of individuals diagnosed with any dementia between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality, PILA, utilizing the healthcare system at least once between 2020 and 2022.

Figure 9: Diagnosis Co-occurrence Networks - Main Comparison



Notes: The figures display diagnosis co-occurrence networks for the year preceding any dementia (ADRD/NND) diagnosis (left panel) and matched controls (right panel). Nodes represent ICD-10 diagnostic categories, with node size proportional to prevalence (the percentage of patients with at least one diagnosis in that category, calculated as the number of unique patients with the diagnosis divided by total patients). Edges connect two categories if they co-occur in the same patient, with edge width indicating the number of patients who share both diagnoses. Node position is determined by a force-directed layout algorithm using closeness and eigenvector centrality measures. Closeness centrality measures how close a diagnostic category is to all other categories in the network (calculated as $(n-1)/\sum_{j\neq c} d(c,j)$, where $d(c,j)$ is the shortest path distance between categories). Eigenvector centrality measures the importance of a category based on its connections to other important categories. The left panel excludes ADRD and NND diagnostic codes themselves, showing only comorbid conditions. Networks exclude generic diagnostic categories (Health Services, Symptoms/Signs, and Special Purposes codes).

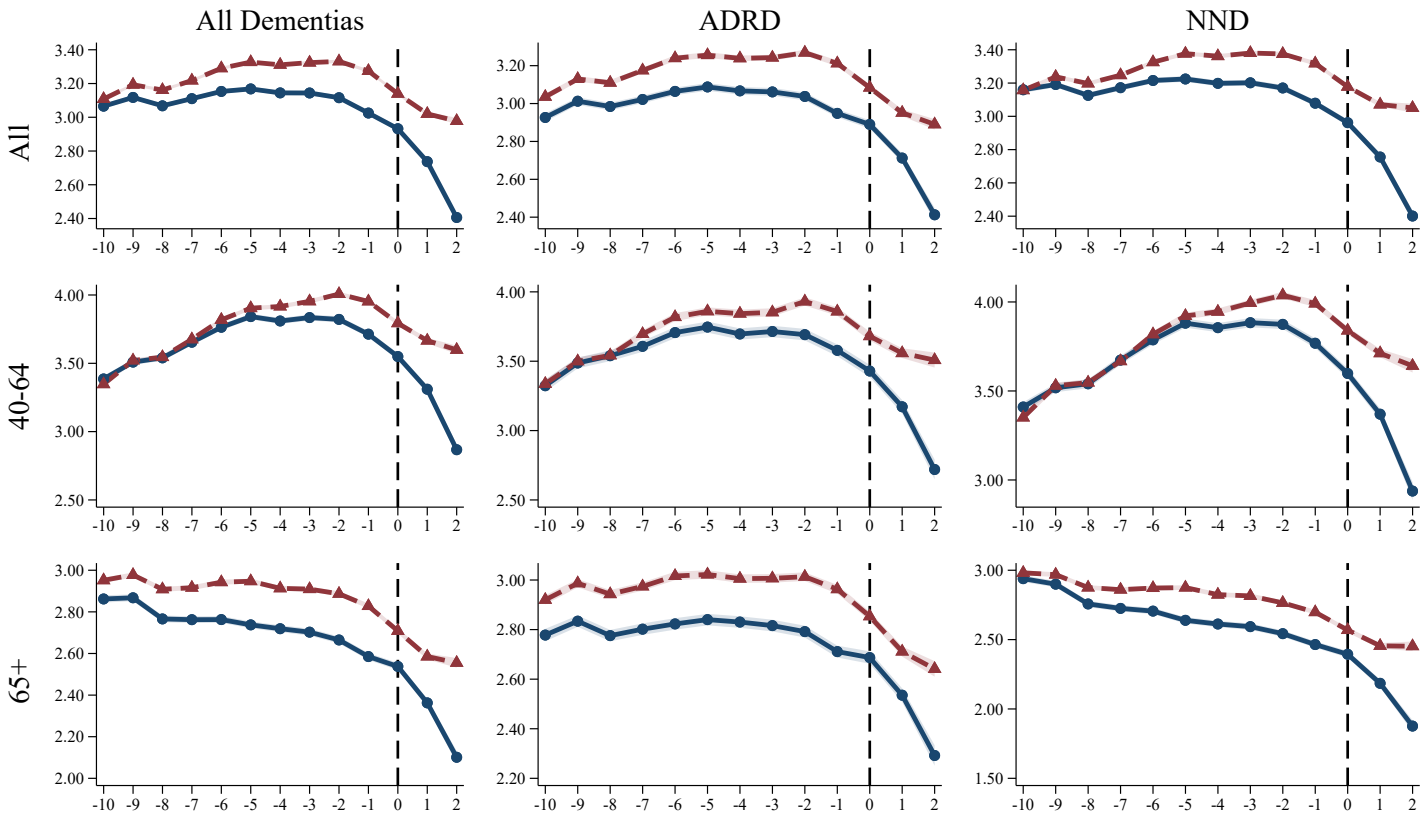
Incidence - All Dementias

Table 4: Diagnosis Category Prevalence, Intensity, and Centrality: Treated vs Controls

	Treated			Control			Difference		
Category	Prevalence	Intensity	Eigen	Prevalence	Intensity	Eigen	Δ Prevalence	Δ Intensity	Δ Eigen
Circulatory	71.77	14.94	0.142	60.86	9.80	0.259	10.91	5.14	-0.117
Endocrine/Metabolic	55.87	8.26	1.000	43.00	5.84	1.000	12.88	2.42	0.000
Genitourinary	43.22	7.76	0.653	32.16	5.78	0.503	11.06	1.98	0.150
Mental Health	41.59	7.25	0.883	10.42	5.70	0.268	31.17	1.55	0.615
Musculoskeletal	40.12	6.01	0.410	32.14	4.42	0.686	7.97	1.59	-0.276
Digestive	36.03	6.19	0.674	29.29	4.65	0.299	6.74	1.54	0.375
Nervous System	29.45	5.64	0.529	11.63	4.00	0.704	17.82	1.64	-0.175
Respiratory	26.14	8.11	0.488	19.21	5.59	0.449	6.93	2.52	0.039
Eye/Adnexa	24.85	3.45	0.271	22.31	3.26	0.871	2.54	0.19	-0.600
Injury/Poisoning	21.62	9.33	0.746	12.14	5.48	0.626	9.48	3.85	0.120
Infectious Diseases	16.26	7.77	0.159	10.41	3.81	0.232	5.85	3.96	-0.073
Skin/Subcutaneous	14.68	4.62	0.708	10.45	3.46	0.218	4.23	1.16	0.490
Ear/Mastoid	13.64	3.23	0.258	9.19	2.73	0.258	4.44	0.50	0.000
Neoplasms	12.97	10.73	0.302	8.80	7.72	0.288	4.17	3.01	0.014
External Causes	7.64	4.72	0.320	4.51	3.63	0.122	3.13	1.09	0.198
Blood Disorders	6.98	5.71	0.468	3.28	4.35	0.027	3.70	1.36	0.441
Congenital	1.84	3.62	0.040	0.96	2.86	0.091	0.88	0.76	-0.051
Pregnancy/Birth	0.50	9.46	0.002	0.30	7.34	0.002	0.21	2.12	0.000
Perinatal	0.12	5.76	0.010	0.09	5.08	0.005	0.03	0.68	0.005

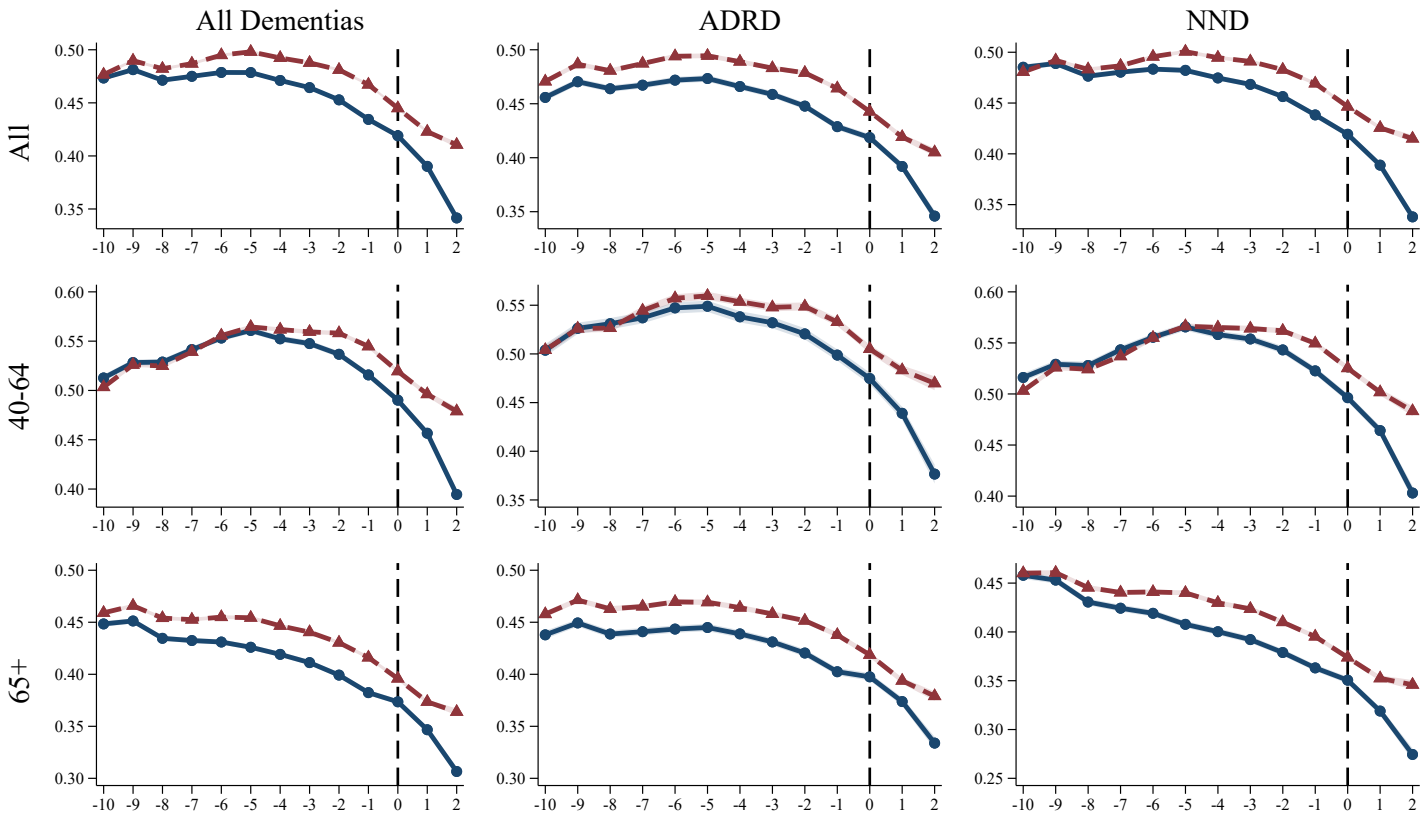
Notes: Prev = Prevalence (% of patients with ≥1 diagnosis); Intens = Average diagnoses per affected patient; Eigen = Eigenvector centrality. Δ = Difference between treated and control groups. Eigenvector centrality is normalized. Sample period: 12 months before ADRD diagnosis.

Figure 10: Labor Market Outcomes of patients and controls: Wages



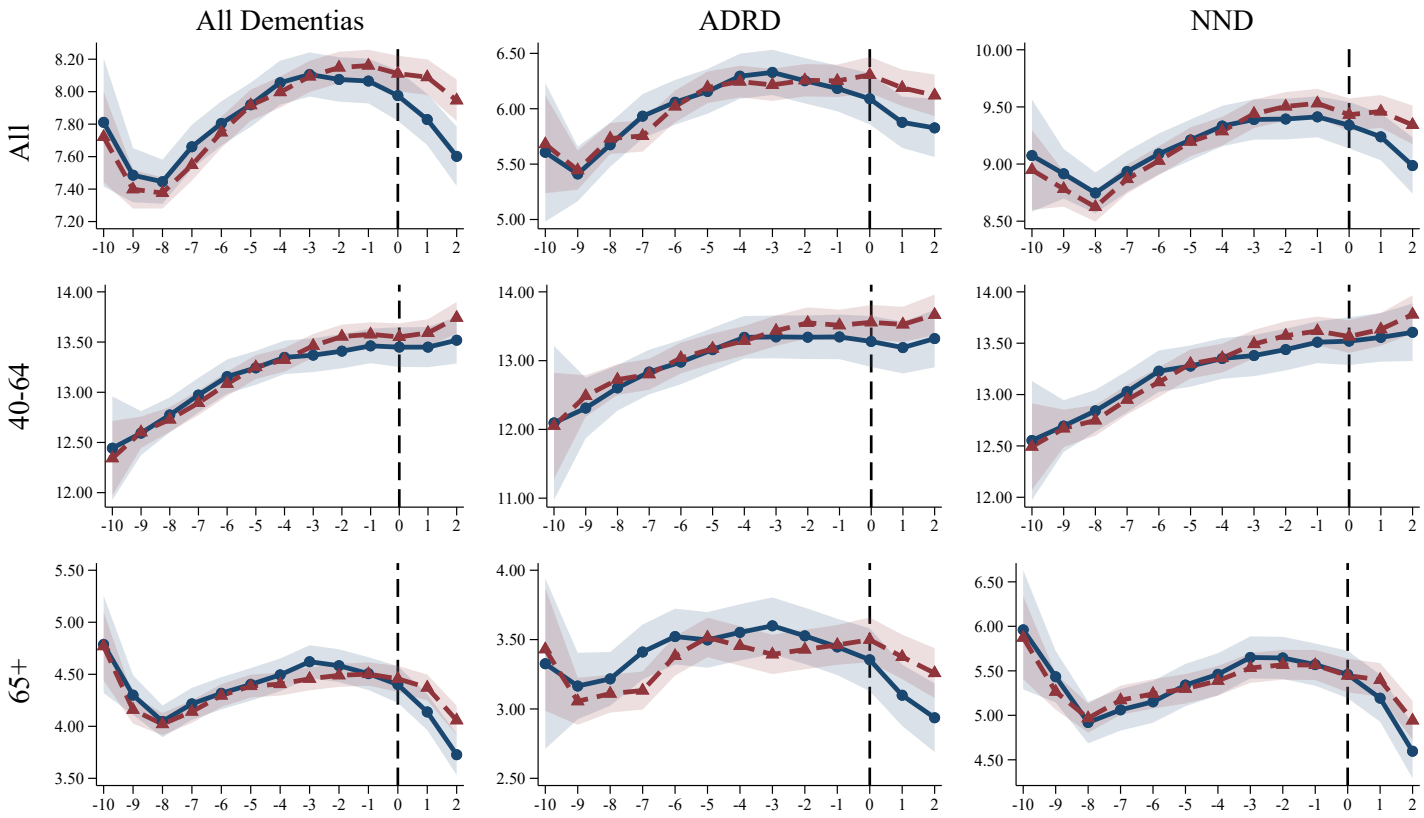
Notes: The figure displays average annual log formal wages (in Colombian pesos) of individuals diagnosed with Alzheimer's Disease and Related Dementias (ADRD) between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality, and appearing on the labor market at least once between 2009 and 2022.

Figure 11: Labor Market Outcomes of patients and controls: Labor Supply



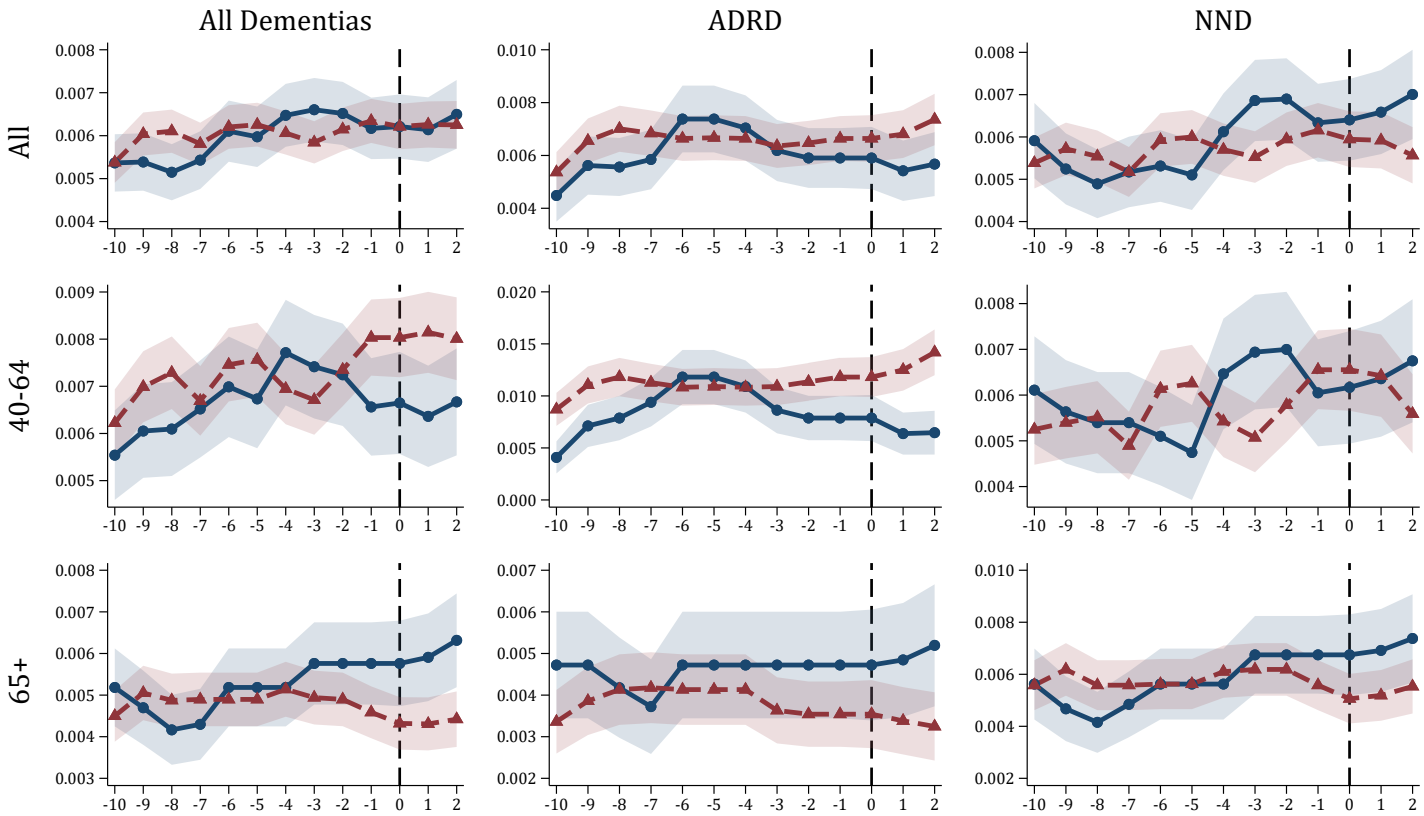
Notes: The figure displays average annual formal employment status of individuals diagnosed with Alzheimer's Disease and Related Dementias (ADRD) between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality, and appearing on the labor market at least once between 2009 and 2022.

Figure 12: Credit Outcomes of patients and controls: Log of Total Balance



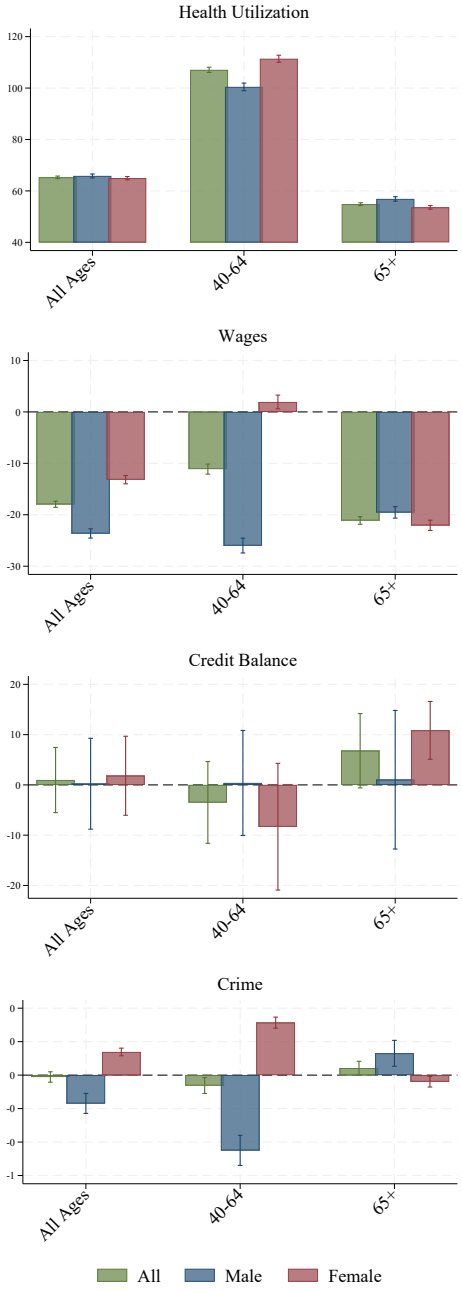
Notes: The figure displays average annual log credit balance (in Colombian pesos) of individuals diagnosed with any dementia between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality.

Figure 13: Crime Outcomes of patients and controls: Conviction duration



Notes: The figure displays average annual conviction duration of individuals diagnosed with Alzheimer's Disease and Related Dementias (ADRD) between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality.

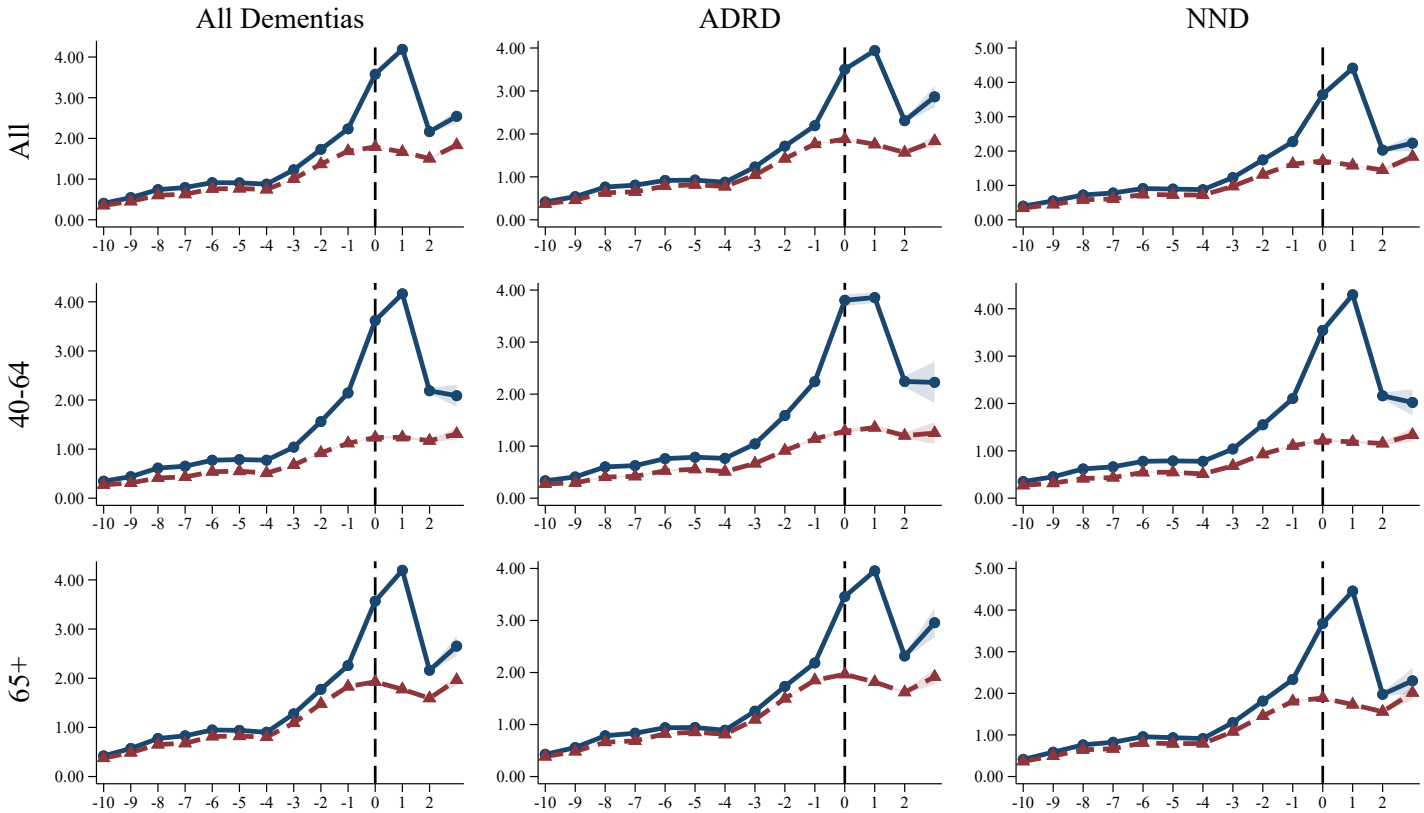
Figure 14: Cumulative loses and extra expenditure: All patients



Notes: The figure displays pre-diagnosis cumulative losses across labor market, health and credit outcomes by gender and age group.

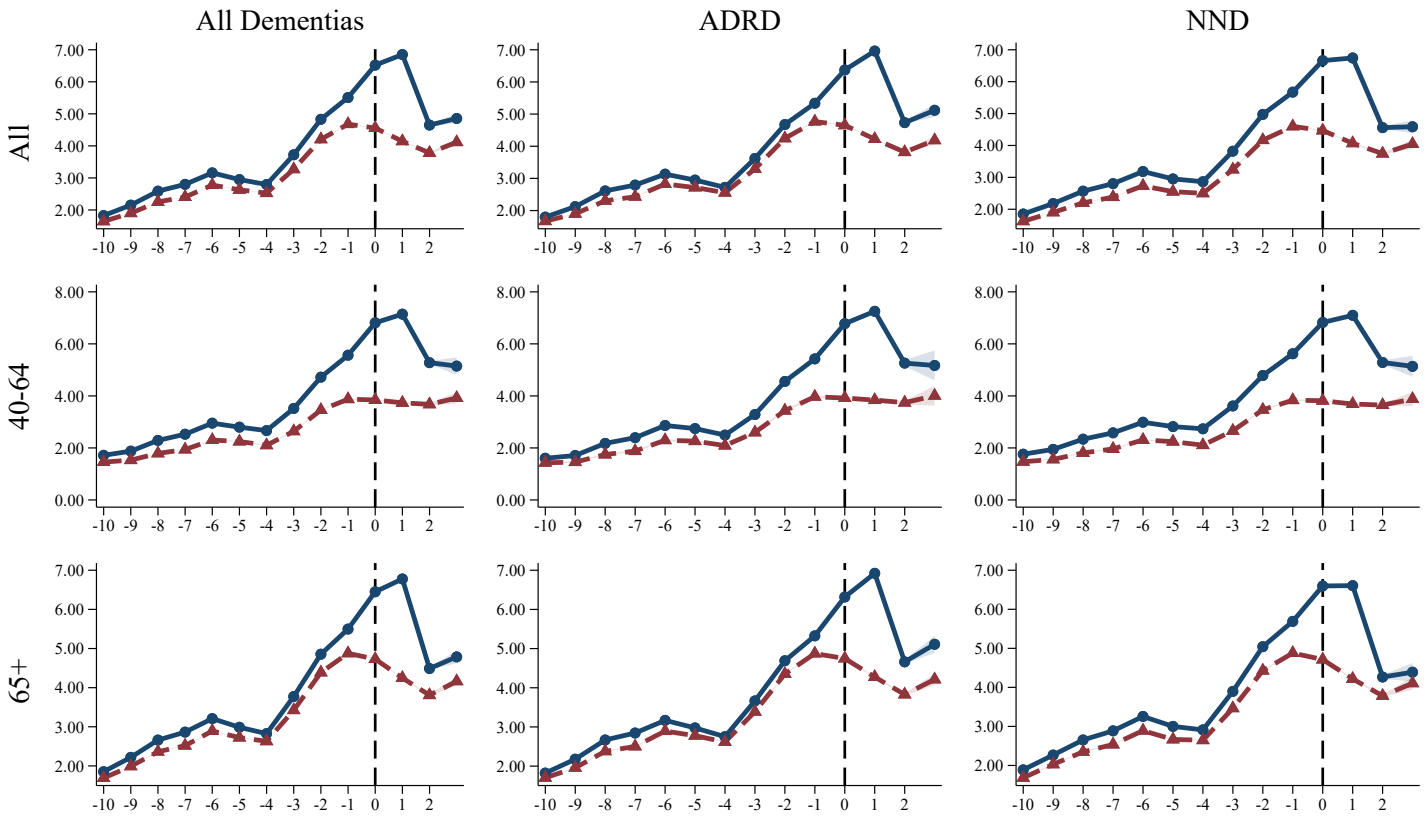
A Additional Tables and Figures

Figure A.1: Health Outcomes of patients and controls: Total Health Events



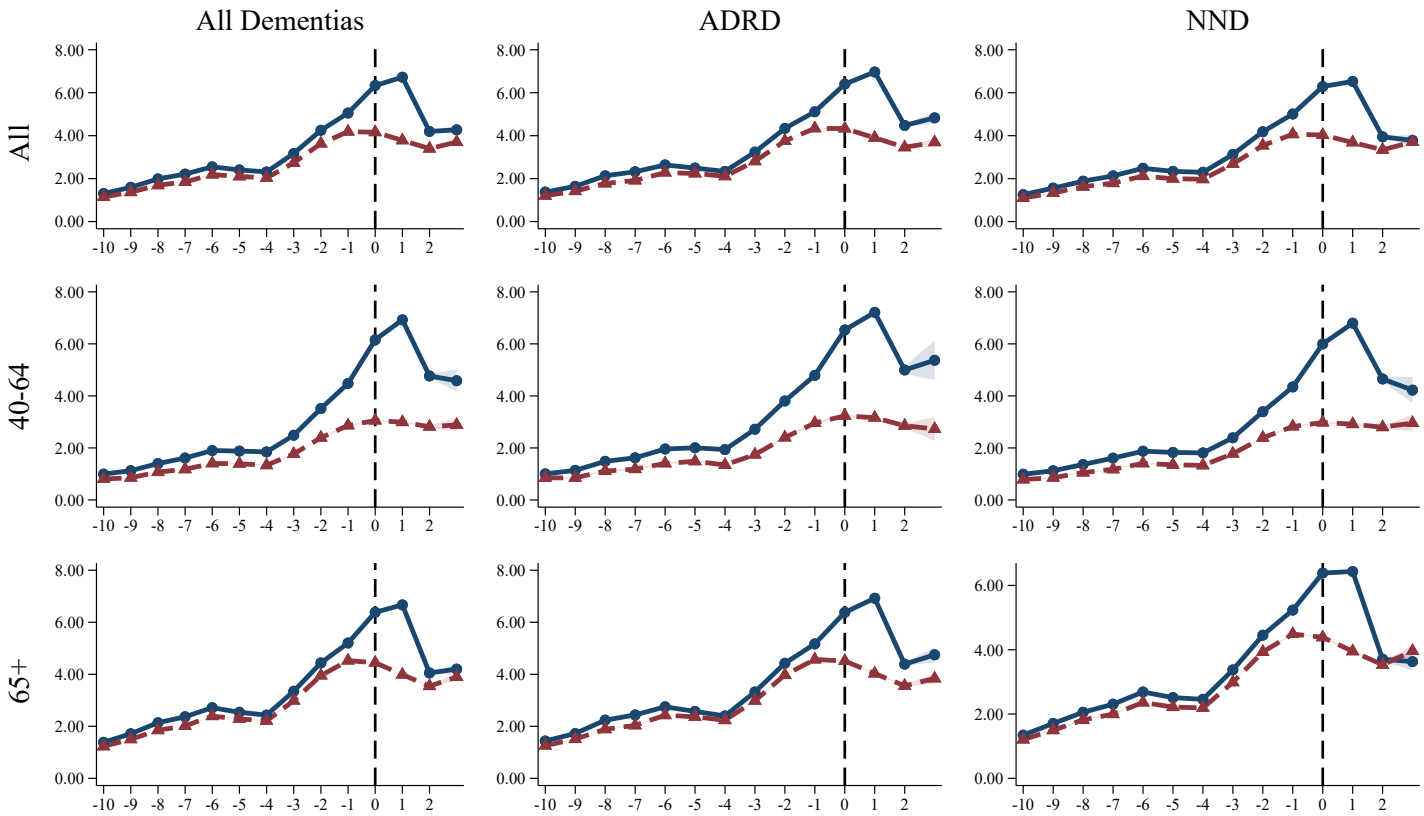
Notes: The figure displays differences in average total health events (consultations, procedures, ER visits, and hospitalizations) of individuals diagnosed with any dementia between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality, and utilizing the healthcare system at least once between 2020 and 2022.

Figure A.2: Health Outcomes of patients and controls: Total Utilization - Women



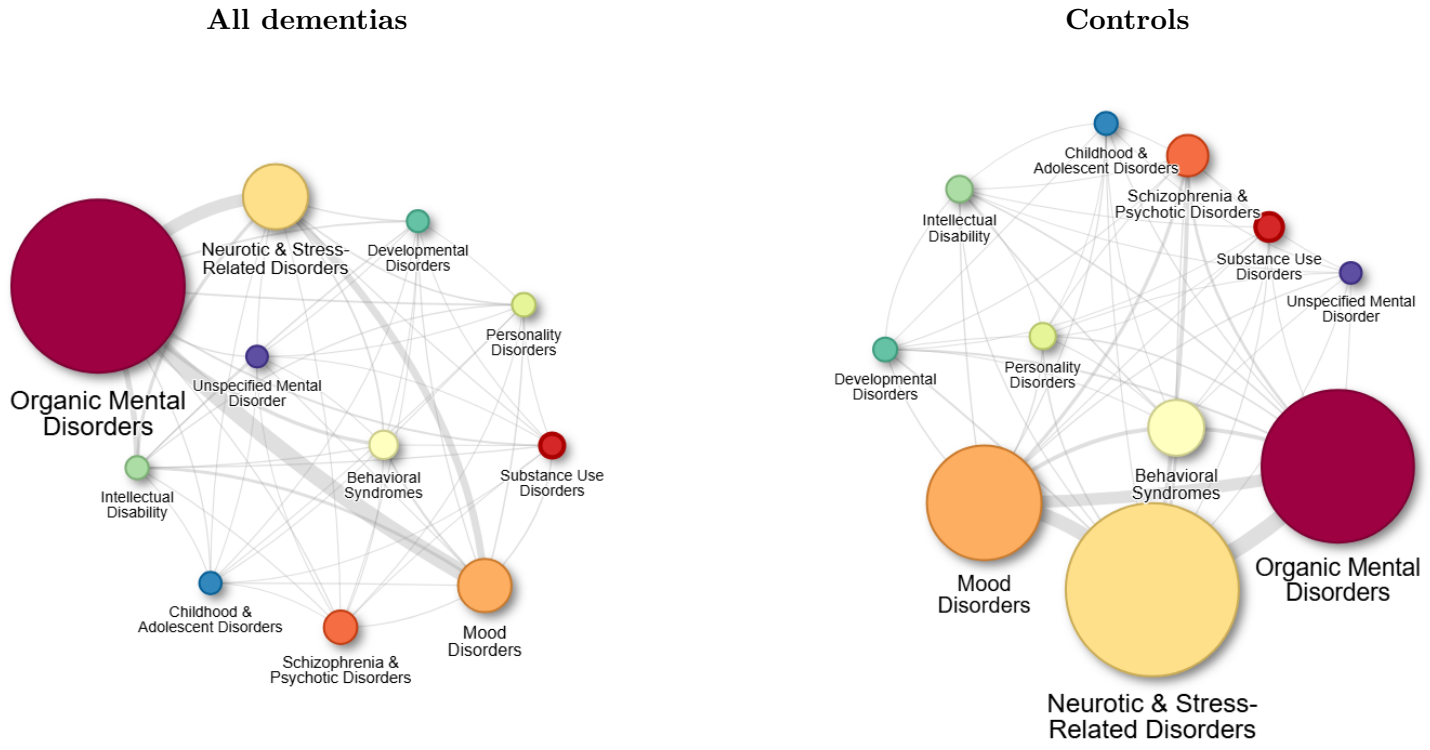
Notes: The figure displays log total utilization (in Colombian pesos) of female individuals diagnosed with any dementia between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality, and utilizing the healthcare system at least once between 2020 and 2022.

Figure A.3: Health Outcomes of patients and controls: Total Utilization - Men



Notes: The figure displays log total utilization (in Colombian pesos) of male individuals diagnosed with any dementia between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality, and utilizing the healthcare system at least once between 2020 and 2022.

Figure A.4: Diagnosis Co-occurrence Networks - Main Comparison



Notes: This figure shows diagnosis co-occurrence networks in the year before ADRD or NND diagnosis (or synthetic diagnosis date for controls). Node size represents incidence of diagnosis category; edge width represents the number of shared patients between diagnosis categories. Centrality measures (degree, betweenness, closeness, eigenvector) are calculated for each diagnostic category. The left panel includes all diagnoses for ADRD and Non-ND patients; the right panel shows the matched control group.

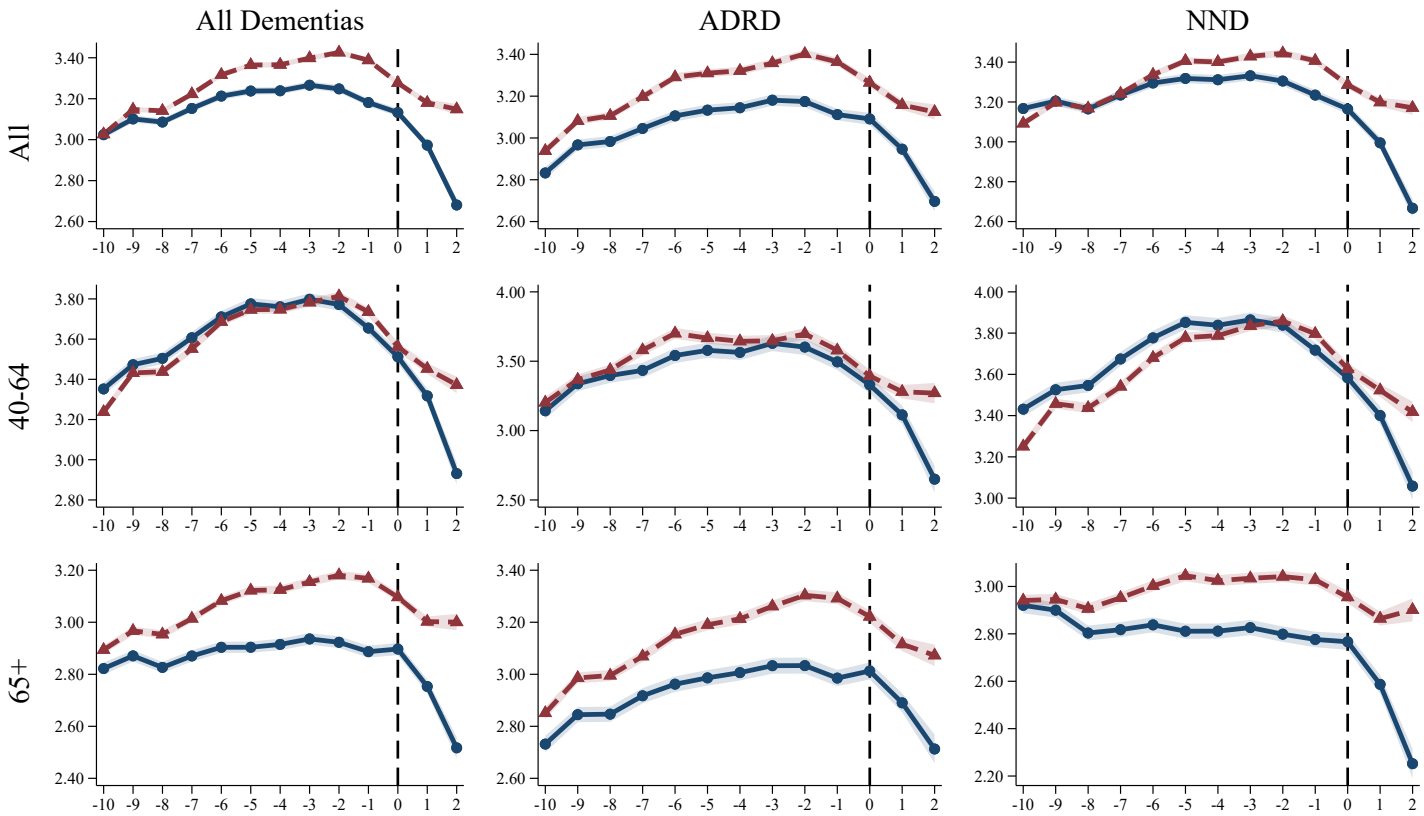
Incidence - All dementias

Table A.1: ICD10 Chapter F Mental Health: Prevalence, Intensity, and Network Centrality - All dementias

	Treated			Control			Difference		
F Block	Prevalence	Intensity	Eigen	Prevalence	Intensity	Eigen	ΔPrevalence	ΔIntensity	ΔEigen
F00–F09 Organic	90.70	3.50	1.000	38.18	7.13	0.959	52.52	-3.63	0.041
F40–F48 Neurotic	26.26	3.82	0.071	44.23	2.80	0.274	-17.97	1.02	-0.203
F30–F39 Mood	19.31	5.19	0.022	27.22	3.21	0.251	-7.91	1.98	-0.229
F20–F29 Schizophrenia	7.32	7.91	0.864	5.83	6.04	0.903	1.49	1.87	-0.039
F50–F59 Behavioral	4.37	2.41	0.176	10.16	2.71	0.053	-5.79	-0.30	0.123
F10–F19 Addictions	**1.70**	**9.53**	**0.746**	**1.86**	**3.49**	**1.000**	**-0.16**	**6.04**	**-0.254**
F60–F69 Personality	1.49	3.05	0.335	1.46	1.69	0.017	0.03	1.36	0.318
F70–F79 Retardation	1.30	7.16	0.080	1.53	4.96	0.077	-0.24	2.20	0.003
F90–F98 Childhood	0.73	2.64	0.059	0.56	1.67	0.024	0.17	0.97	0.035
F80–F89 Development	0.57	3.23	0.016	0.84	6.96	0.058	-0.27	-3.73	-0.042
F99 Unspecified	0.36	2.66	0.032	0.14	1.72	0.009	0.21	0.94	0.023

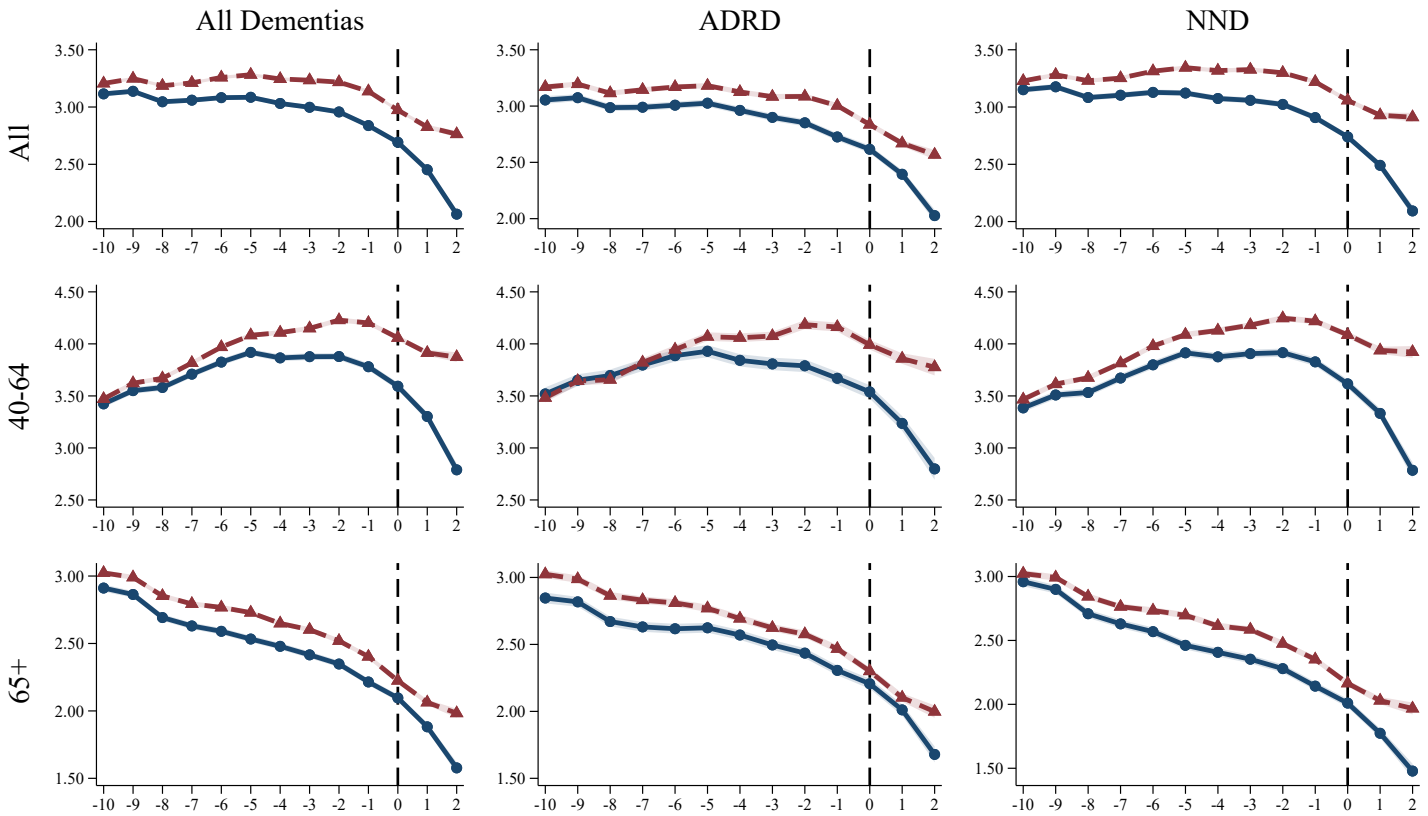
Notes: Prev = prevalence (% of patients). Intensity = average diagnoses per affected patient. Eigen = eigenvector centrality. Δ = treated minus control. Centrality measures normalized. Period: 12 months prior to any dementia diagnosis. Bold row indicates Substance Use Disorders (F10–F19).

Figure A.5: Labor Market Outcomes of patients and controls: Wages - Women



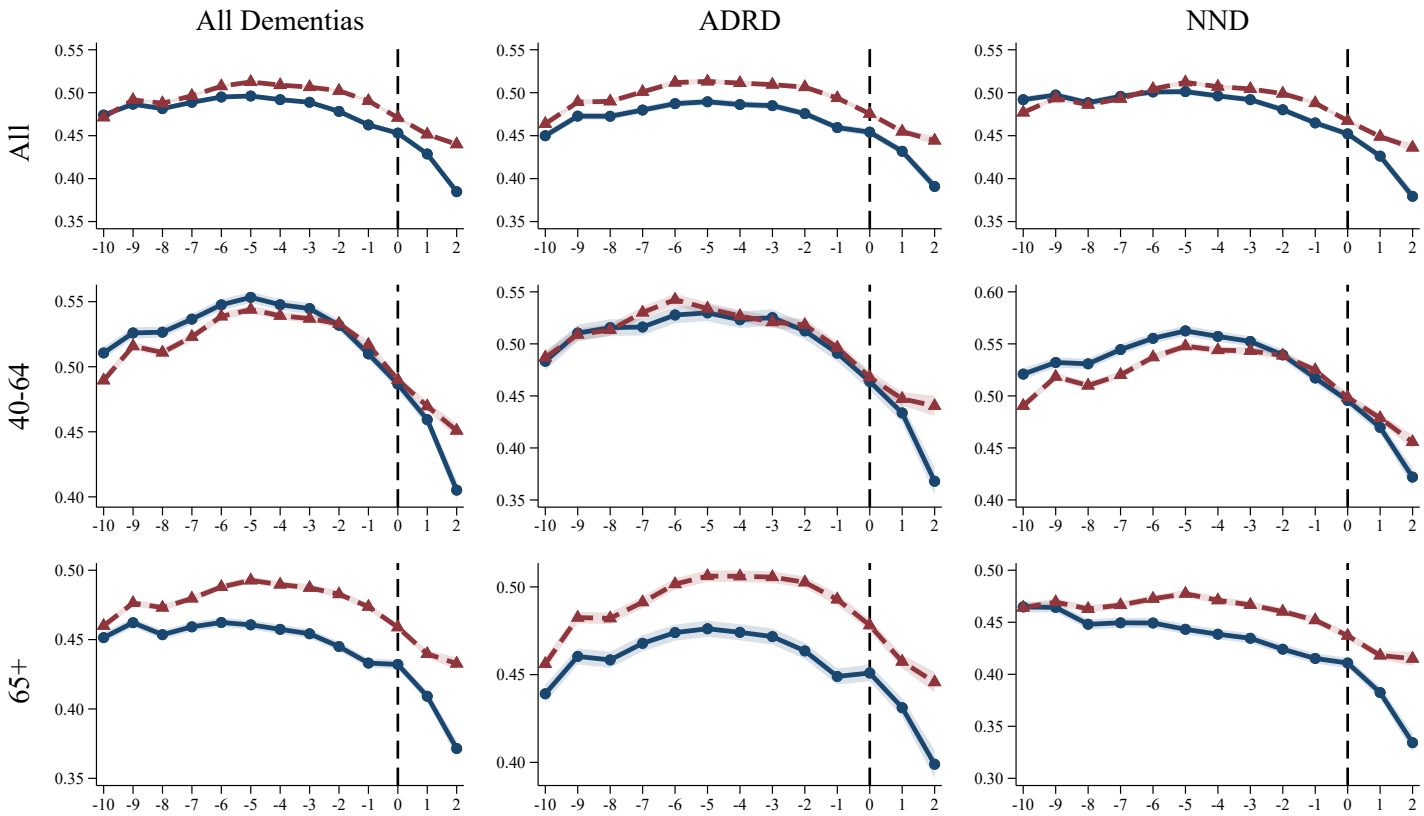
Notes: The figure displays average annual log formal wages (in Colombian pesos) of female individuals diagnosed with any dementia between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality, and appearing on the labor market at least once between 2009 and 2022.

Figure A.6: Labor Market Outcomes of patients and controls: Wages - Men



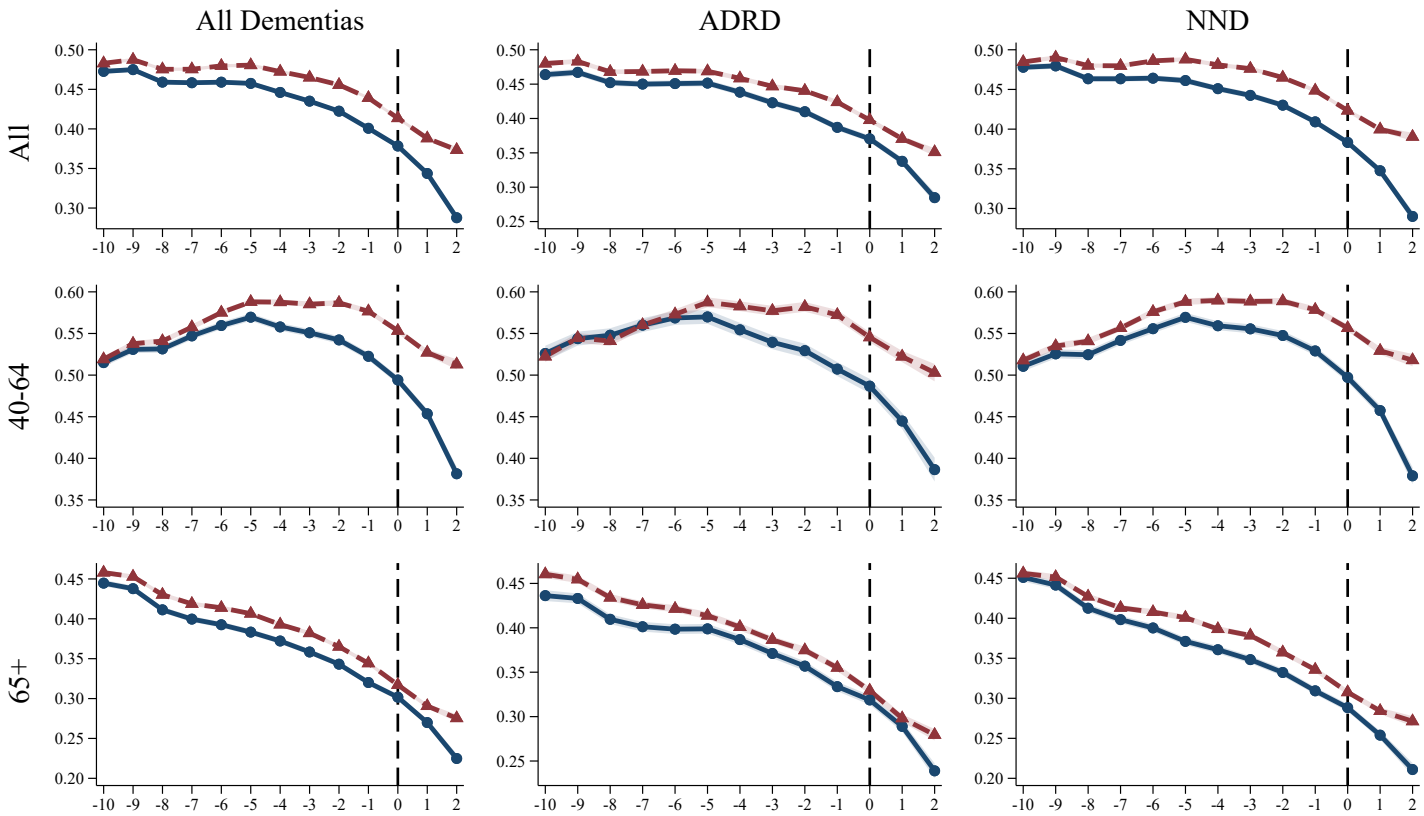
Notes: The figure displays average annual log formal wages (in Colombian pesos) of male individuals diagnosed with any dementia between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality, and appearing on the labor market at least once between 2009 and 2022.

Figure A.7: Labor Market Outcomes of ADRD patients and controls: Labor Supply - Women



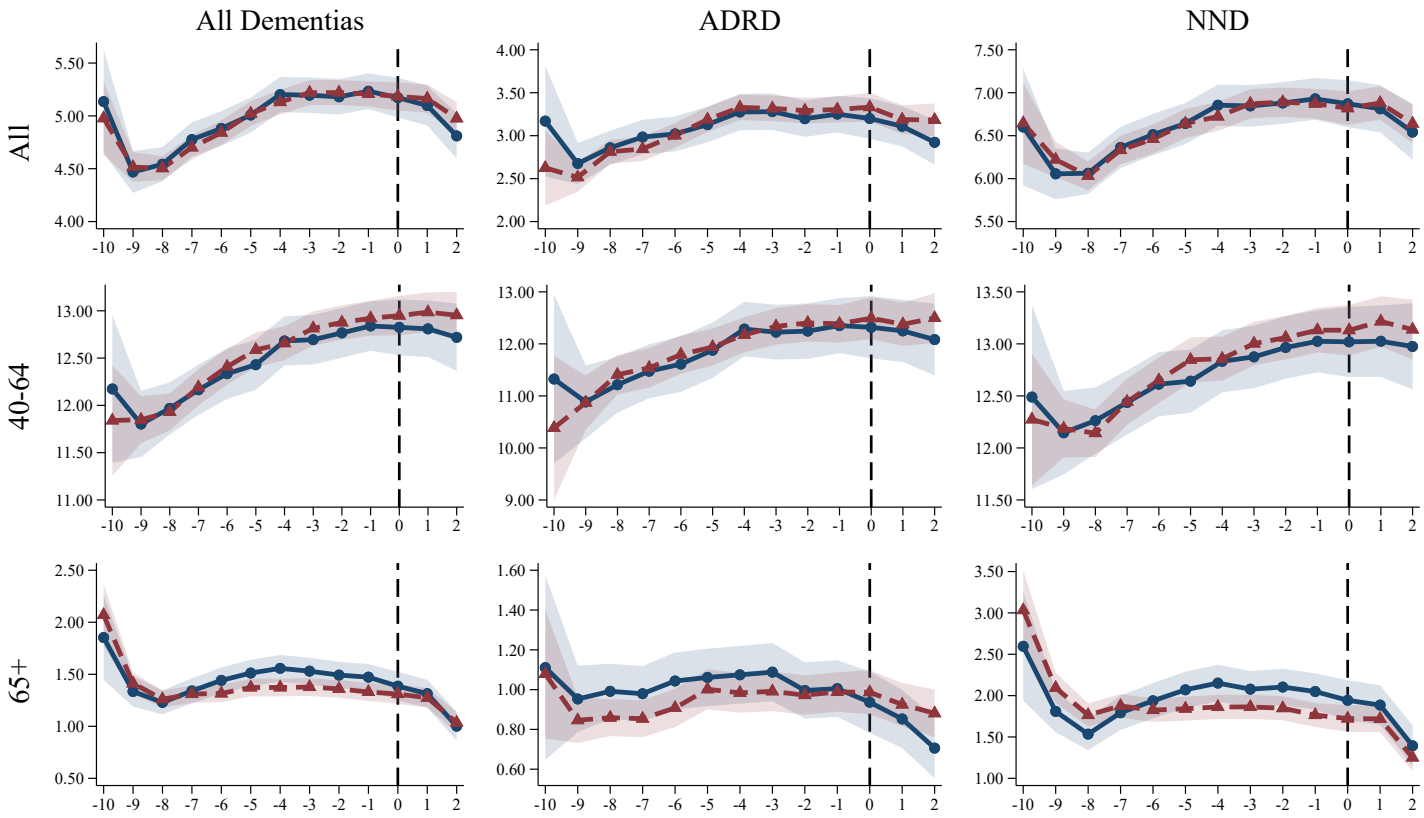
Notes: The figure displays labor market outcomes for individuals diagnosed with any dementia between 2020 and 2022 (treated group), and a matched control group without an any dementia diagnosis during the observation window. The figure shows the average annual formal employment status. Year 0 corresponds to the year of first diagnosis. Averages are computed separately by year relative to diagnosis. Treated individuals (shown in navy) and control individuals (shown in maroon) are matched on age, sex, and municipality..

Figure A.8: Labor Market Outcomes of ADRD patients and controls: Labor Supply - Men



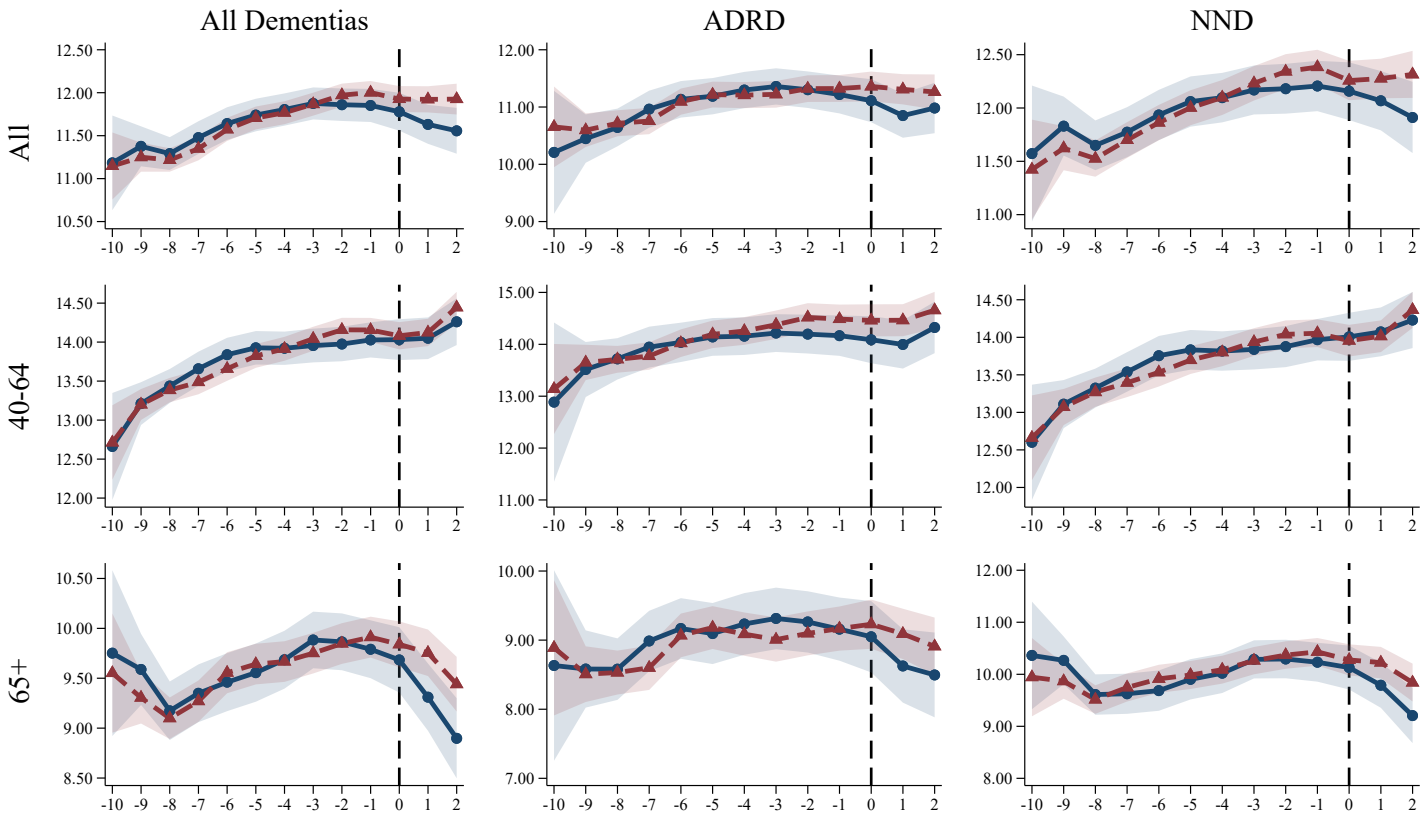
Notes: The figure displays labor market outcomes for individuals diagnosed with any dementia between 2020 and 2022 (treated group), and a matched control group without an ADRD diagnosis during the observation window. The figure shows the average annual formal employment status. Year 0 corresponds to the year of first diagnosis. Averages are computed separately by year relative to diagnosis. Treated individuals (shown in navy) and control individuals (shown in maroon) are matched on age, sex, and municipality..

Figure A.9: Credit Outcomes of patients and controls: Log of Total Balance - Women



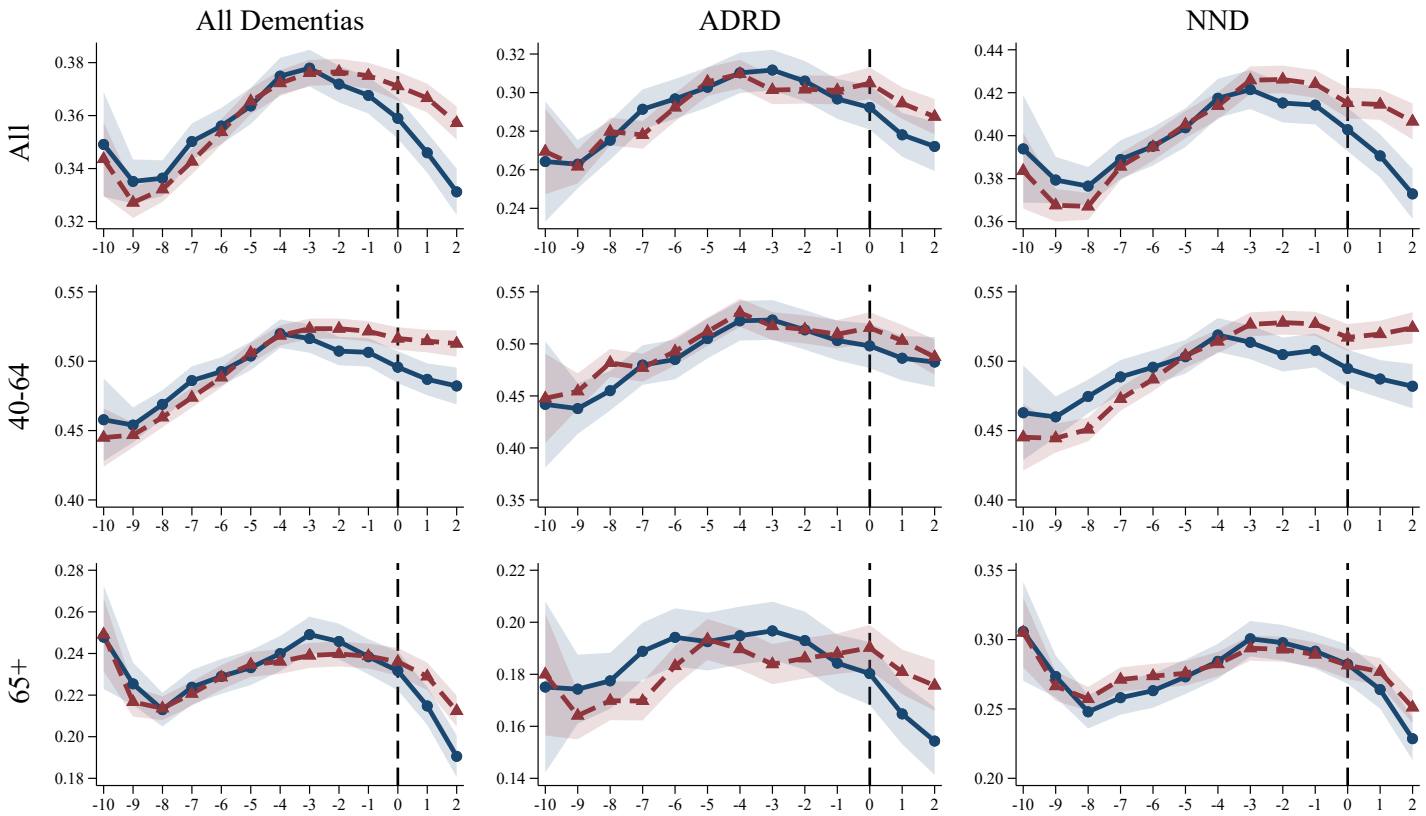
Notes: The figure displays average annual log credit balance (in Colombian pesos) of female individuals diagnosed with any dementia between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality.

Figure A.10: Credit Outcomes of patients and controls: Log of Total Balance - Men



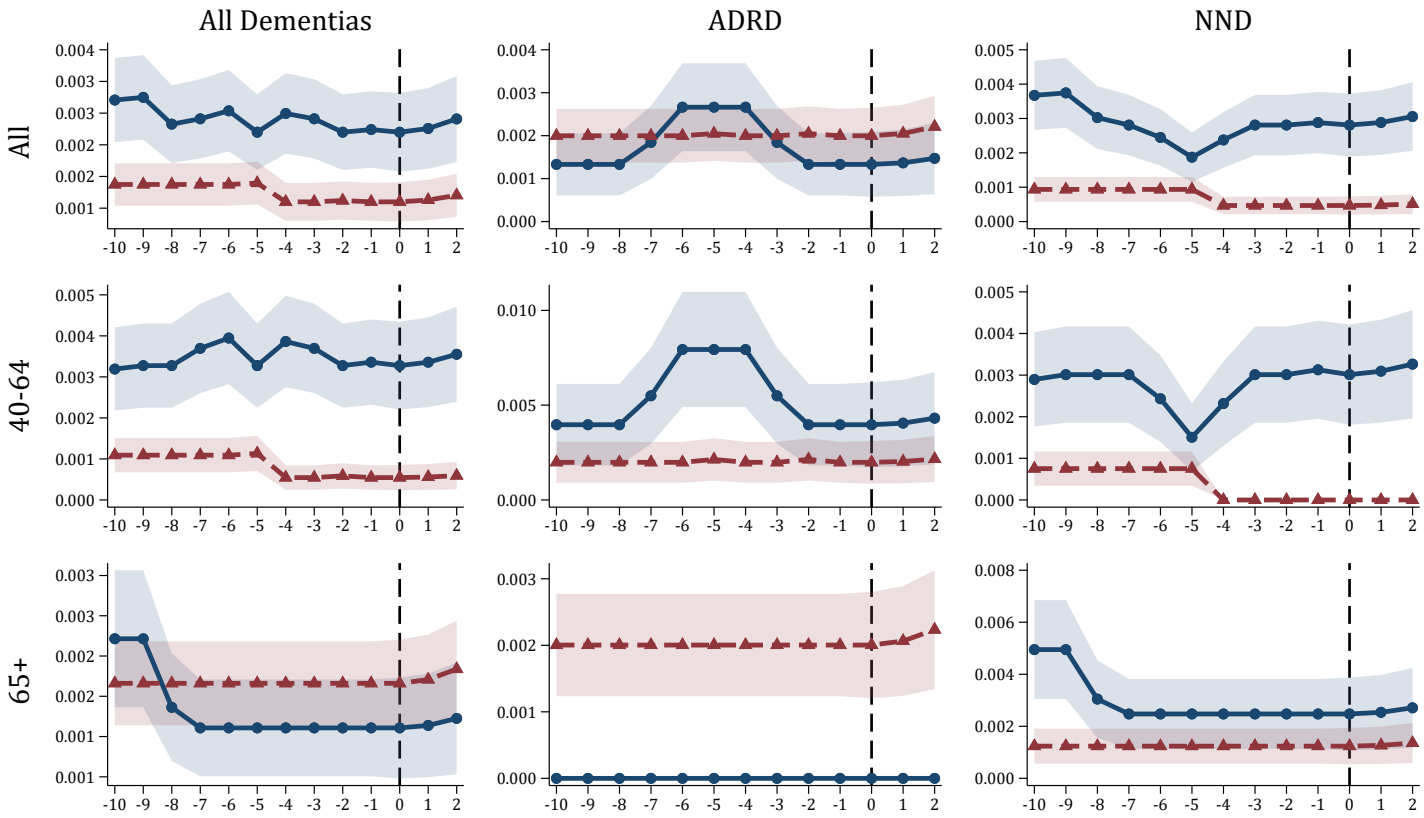
Notes: The figure displays average annual log credit balance (in Colombian pesos) of male individuals diagnosed with any dementia between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality.

Figure A.11: Credit Outcomes of patients and controls: Total Credits



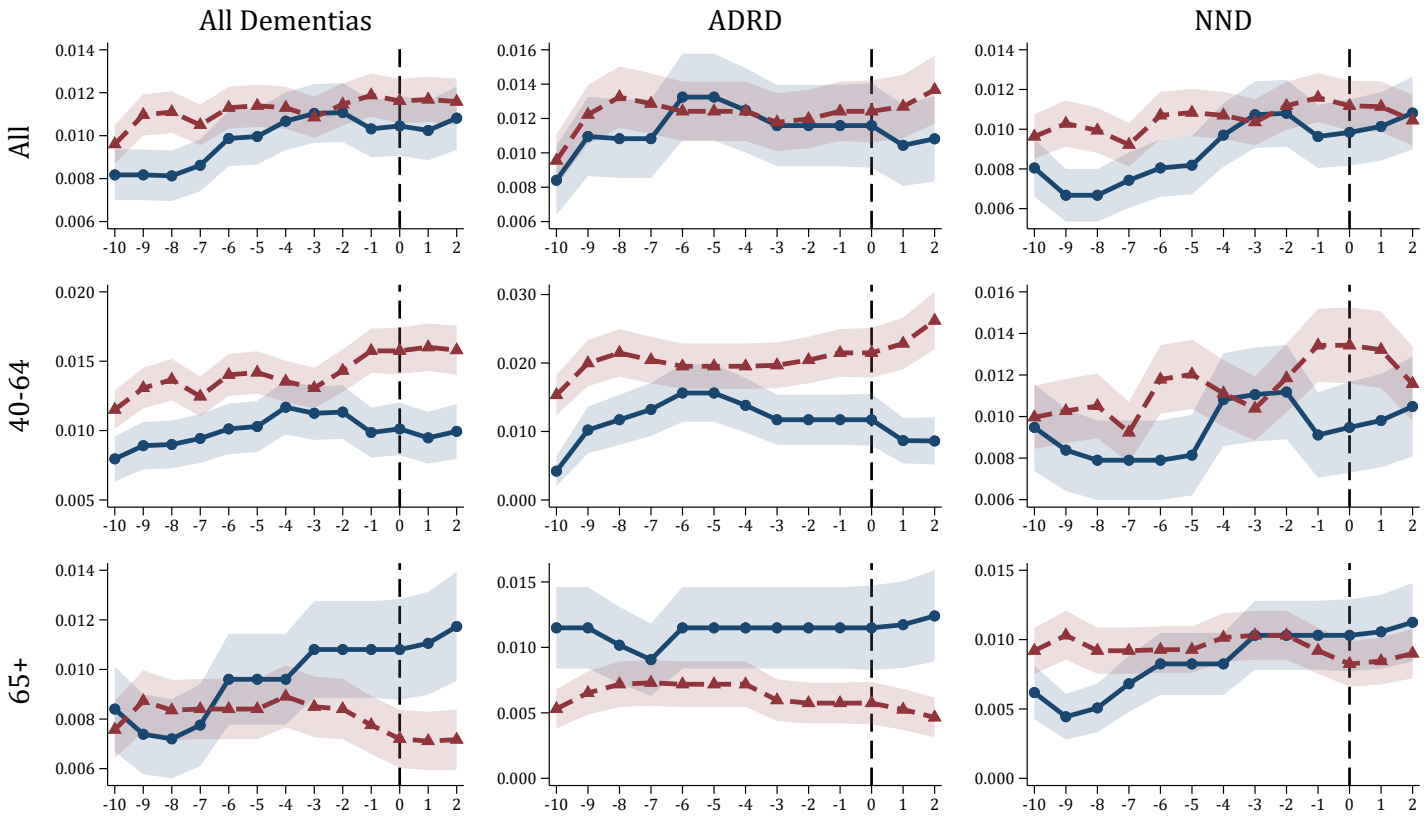
Notes: The figure displays annual average number of credits of individuals diagnosed with any dementia between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality.

Figure A.12: Crime Outcomes of patients and controls: Crimes - Women



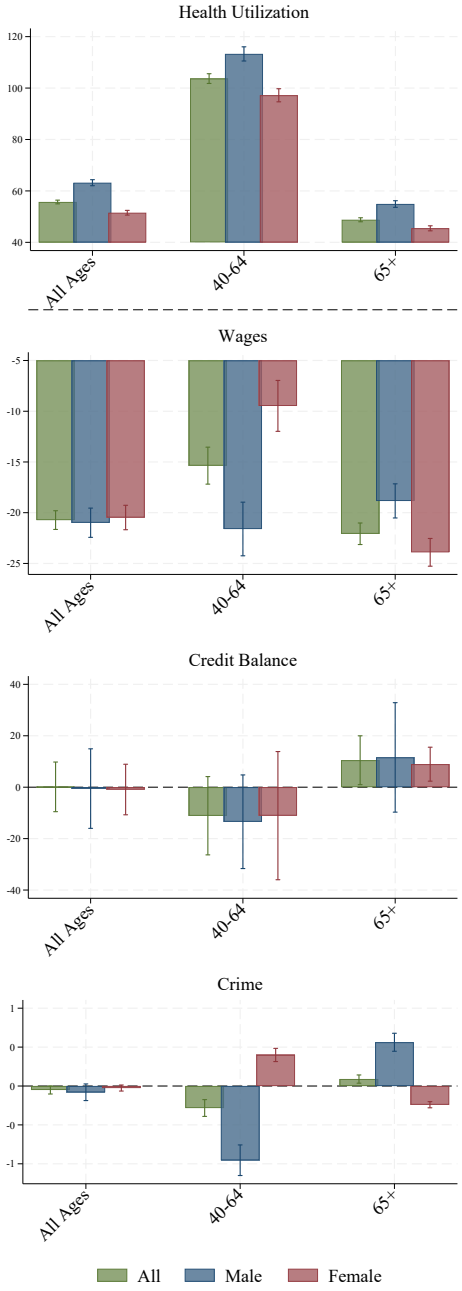
Notes: The figure displays average annual conviction duration of female individuals diagnosed with any dementia between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality.

Figure A.13: Crime Outcomes of patients and controls: Crimes - Men



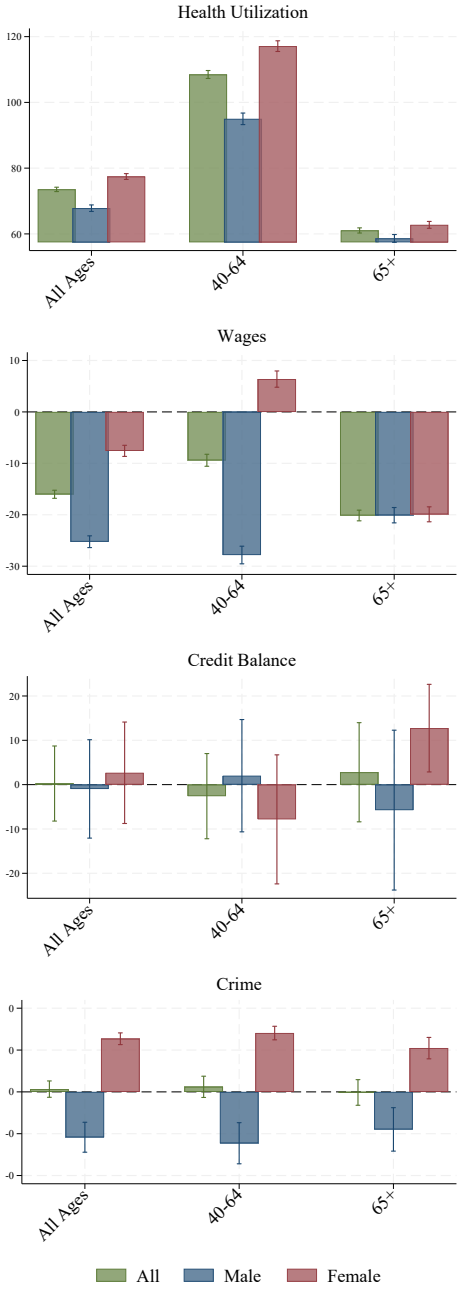
Notes: The figure displays average annual conviction duration of male individuals diagnosed with any dementia between 2020 and 2022 (navy), and a matched control group (maroon). The x axis reflect the years relative to the first diagnosis. Treated and control individuals are matched on age, sex, municipality.

Figure A.14: Cumulative loses and extra expenditure: ADRD Patients



Notes: The figure displays pre-diagnosis cumulative losses across labor market, health and credit outcomes by gender and age group.

Figure A.15: Cumulative loses and extra expenditure: Non-ND Patients



Notes: The figure displays pre-diagnosis cumulative losses across labor market, health and credit outcomes by gender and age group.