NBER WORKING PAPER SERIES

PATIENT PEER EFFECTS: EVIDENCE FROM NURSING HOME ROOM ASSIGNMENTS

Alden Cheng
Martin B. Hackmann

Working Paper 34538
http://www.nber.org/papers/w34538

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2025

We thank Yiqun Chen, Ashvin Gandhi, Erzo Luttmer, Devin Pope, Tamara Konetzka, as well as participants of the Midwest Econometrics Group Conference and Occasional California Health Economics Conference for helpful comments and suggestions. Funding from the National Institute on Aging grant #P01 AG005842-29 is gratefully acknowledged. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Patient Peer Effects: Evidence from Nursing Home Room Assignments
Alden Cheng and Martin B. Hackmann
NBER Working Paper No. 34538
December 2025
JEL No. D62, I11, I12, I18, J14

ABSTRACT

We provide causal evidence that patient peer effects generate mortality impacts comparable to provider quality differences. Drawing on administrative records covering 2.6 million stays (2000–2010) across 7,200 U.S. nursing homes, we exploit plausibly exogenous roommate assignments identified through unique room identifiers. We estimate that assignment to a roommate diagnosed with Alzheimer's disease (AD) or Alzheimer's disease related dementias (ADRD), relative to placement in a private room, increases 90-day mortality by 2.1 percentage points (14% of baseline)—equivalent to receiving care at a nursing home one full standard deviation worse in quality. Effects differ sharply by patient type: patients with AD/ADRD benefit substantially from cognitively healthy roommates but not from private rooms, suggesting important peer monitoring and support roles. In contrast, mortality of patients without AD/ADRD does not depend on roommate cognitive health but is reduced in private rooms. A simple assignment rule exploiting this heterogeneity could reduce overall mortality by 0.8 percentage points without additional resources.

Alden Cheng
National Bureau of Economic Research
and University of Illinois, Urbana-Champaign
alden15@nber.org

Martin B. Hackmann
University of California, Los Angeles
Department of Economics
and NBER
mbhackmann@gmail.com

1 Introduction

With declining fertility rates, rising childlessness, and increasing prevalence of Alzheimer's disease (AD) and Alzheimer's disease related dementias (ADRD) in the population, nursing homes have become an increasingly important part of the healthcare sector. Nearly 1.2 million Americans reside in nursing homes, supported by $91 billion in annual Medicaid and Medicare expenditures (Kaiser Family Foundation, 2025), and 56% of individuals currently aged 57–61 are projected to stay in a nursing home in the future (Hurd, Michaud, and Rohwedder, 2017). Accordingly, a growing body of work in economics studies the production of health in nursing homes, and incentives driving facilities' quality decisions (Hackmann, 2019; Gandhi et al., 2024; Einav, Finkelstein, and Mahoney, 2025). Yet a fundamental question has received surprisingly little attention: does your nursing home roommate affect your health? Nearly 80% of patients are assigned to shared rooms upon admission, yet beyond well-documented infection transmission (Brown et al., 2021; Konetzka, Grabowski, and Mor, 2024), we know very little about whether—and how—roommates influence each other's health.

This paper provides the first causal evidence that patient peer effects, operating largely through non-infectious channels, generate mortality impacts comparable to provider quality differences. Assignment to a roommate with AD/ADRD, relative to placement in a private room, increases 90-day mortality by 2.1 percentage points (14% of baseline)—equivalent to receiving care at a nursing home one full standard deviation worse in quality. Moreover, these peer effects vary significantly by observable patient characteristics, and a simple room assignment rule may reduce overall mortality rates by 0.8 percentage points relative to the status quo, without additional staff or facility modifications. These findings reveal profound asymmetries in how patients interact as productive inputs, with immediate implications for care delivery, facility design, and $91 billion in annual public expenditures.

Understanding patient peer effects matters particularly because of what we do—and do not—know about quality variation in patient care. A substantial literature documents large mortality and quality differences between hospitals (Geweke, Gowrisankaran, and Town, 2003; Chandra et al., 2016a; Doyle, Graves, and Gruber, 2019; Hull, 2018) and nursing homes (Olenski and Sacher, 2024; Einav, Finkelstein, and Mahoney, 2025; Cheng, 2023). Yet we know relatively little about what drives variation within facilities—whether patients receiving care from identical staff, in the same building, with the same aggregate inputs, experience systematically different outcomes. Recent evidence shows that most Black-white disparities in nursing home quality occur within rather than between facilities (Einav et al., 2025), underscoring that facility-level measures may mask important heterogeneity in patient experiences. Our findings reveal an important mechanism driving such within-facility variation: peer effects generate mortality differences comparable to between-facility quality gaps that have motivated decades of policy interventions such as minimum staffing regulations and inspections. This suggests micro-level assignment policies may be as consequential as macro-level facility choice—a margin entirely absent from current quality measurement, reimbursement policy, and consumer information systems.

The scarcity of credible evidence on patient peer effects is not without reason. Roommate assignment is rarely random; facilities may match roommates by acuity, diagnosis, or behavior. Separating peer effects from selection, common shocks, and reflection problems requires granular assignment data and credible exogenous variation—both scarce in existing research. We overcome these challenges using administrative data covering 2.6 million patient-stays across over 7,200 U.S. nursing facilities (2000–2010). We construct a novel database linking high-frequency patient-to-room assignments and exploit quasi-random variation in roommate characteristics driven by bed availability at admission. Conditional on facility-by-time fixed effects, the types of rooms available at the precise time of admission generates plausibly exogenous variation in whether patients receive private rooms, and if not, their roommate's cognitive health and acuity.

Our empirical analysis yields three main findings. First, patient peer effects are quantitatively large and generate substantial within-facility variation in outcomes. Moving from a private room to one shared with an Alzheimer's patient increases 90-day mortality by 2.1 percentage points (14% relative to baseline). Our findings suggest that these effects operate largely through non-infectious channels: vaccination status explains little of the mortality differences between empty and shared rooms.

Second, we document striking heterogeneity revealing important substitutabilities in health production. Patients with AD/ADRD benefit substantially from cognitively healthy roommates—experiencing a 5.1 percentage point mortality reduction—but do not gain from private rooms. Conversely, patients without AD/ADRD are unharmed by cognitively impaired roommates but do benefit from placement in a private room. Our preferred interpretation is that cognitively healthy patients can provide important monitoring and behavioral support for patients with AD/ADRD, generating a one-directional positive externality. This asymmetry is difficult to explain by infection control, privacy preferences, or symmetric social interaction—it points to peer production of health through caregiving spillovers.

Third, these interaction effects have profound policy implications. Because patients with AD/ADRD benefit substantially from cognitively intact roommates—while cognitively intact patients are unharmed by roommates with AD/ADRD—mixing patients by cognitive status can reduce average mortality. We formalize this insight through simulated assignment rules, showing that mortality-minimizing assignment policies differ markedly from current practice and can yield large reductions in average mortality. While current assignments tend to match patients by cognitive status, the mortality-minimizing policy groups patients with and without AD/ADRD as roommates, leading to reductions in 90-day mortality of at least 0.8 percentage points. Crucially, these gains do not necessarily require additional staffing or capital.

Evidence on mechanisms supports the monitoring interpretation. Peer effects are twice as large in facilities with below-median staffing, suggesting roommates substitute for professional supervision when care is inadequate. Effects are also concentrated among patients not in specialized Alzheimer's units, which provide enhanced monitoring and environmental accommodations. Both patterns indicate peer health substitutes for formal care inputs—precisely what theory predicts if cognitively healthy roommates provide surveillance and behavioral regulation.

We contribute to three strands of literature. First, we extend the peer effects literature which has mostly focused on education and younger populations (Sacerdote, 2001; Carrell, Fullerton, and West, 2009; Abdulkadiroğlu, Angrist, and Pathak, 2014), by providing the first causal evidence that patient peer effects in healthcare among elderly populations—operating largely through non-infectious channels—generate mortality impacts comparable to provider quality differences. Assignment to a roommate with AD/ADRD increases 90-day mortality by 2.1 percentage points (14% of baseline), equivalent to receiving care at a facility one standard deviation worse in quality. Methodologically, while most peer effects studies exploit one-time random assignment, we leverage high-frequency variation in bed availability at admission—a novel identification strategy applicable to other institutional settings where assignment is endogenous but vacancies are quasi-random.

Second, we contribute to the health production function literature. Previous work documents substantial value-added differences between hospitals (Chandra et al., 2016a; Hull, 2018) and nursing homes (Olenski and Sacher, 2024; Einav, Finkelstein, and Mahoney, 2025; Cheng, 2023).[1] We identify a distinct margin: within-facility heterogeneity driven by peer interactions, holding aggregate inputs fixed. We show this margin is quantitatively comparable to between-facility quality differences that have motivated decades of policy interventions. This challenges the conventional focus on aggregate resources and suggests micro-level assignment policies may be as consequential as macro-level facility choice—a margin entirely absent from current quality measurement, reimbursement policy, and consumer information systems.

Third, we contribute to a small but growing literature documenting within-facility variation in nursing home care quality. Prior work has shown racial differences in inputs (medication use, feeding tubes, vaccinations) and outcomes (re-hospitalization, pressure ulcers, home discharges) among patients within the same facility (Einav et al., 2025). We identify a distinct source of within-facility variation: room assignments and peer effects.

Our analysis relates closely to concurrent work by McWilliam (2025), who examines a comprehensive set of outcomes for private versus shared rooms among post-acute nursing home patients, finding no significant improvements in mortality, hospital readmission, discharge to home, length of stay, or resident mood. Our findings complement and extend this work in three ways. First, we examine peer effects within shared rooms, showing that outcomes depend not just on room type but also on roommate characteristics—a margin absent from private-vs-shared comparisons. Second, using a larger sample covering broader populations, we find smaller private-room effects on mortality for post-acute patients (consistent with McWilliam) but substantial effects for long-stay patients with greater peer exposure. Third, we show the "private room effect" is highly heterogeneous: beneficial for cognitively healthy patients, but neutral or even harmful for patients with AD/ADRD who gain more from cognitively healthy roommates.

[1] Prior work also documents provider-patient complementarities from communication frictions and bias (Alsan and Wanamaker, 2018; Alsan, Garrick, and Graziani, 2019; Greenwood et al., 2020).

This paper proceeds as follows. Section 2 describes the nursing home setting and our data. Section 3 outlines our IV strategy, discussing the identifying assumptions and providing evidence that they are satisfied. Section 4 presents our peer effect estimates and explores mechanisms. Section 5 concludes.

2 Background and Data

2.1 Background

Nursing homes represent a substantial component of the U.S. long-term care system. Approximately 1.2 million Americans reside in certified nursing homes at any given time (Kaiser Family Foundation, 2025). Total revenues for U.S. nursing care facilities are roughly $140 billion annually, reflecting the sector's significant economic footprint. Medicare and Medicaid together finance about than two-thirds of nursing home spending, underscoring the sector's central importance for both public health and fiscal sustainability.

The importance of this sector is growing with demographic change. The prevalence of Alzheimer's disease and related dementias (AD/ADRD) is projected to rise from over 6 million Americans today to 13.8 million by 2050 (Matthews et al., 2019). Individuals with dementia are significantly more likely to require nursing home care than those without cognitive impairment; between 2017 and 2019, more than three million nursing home patients have been diagnosed with Alzheimer's disease and related dementias (ADRD) (Mukamel et al., 2023). Because such a large proportion of patients have Alzheimer's disease or related dementias, understanding how facility practices shape outcomes for this population has become an urgent policy priority.

2.1.1 Room Sharing: Benefits, Costs, and Policy Debates

A distinctive feature of nursing home care is that most patients share rooms. In our sample, detailed below, almost 80% of patients are assigned to a shared room upon admission. Unlike hospitals, where room sharing is typically brief, nursing home stays often last months or years, creating extended peer exposure.

Room sharing involves important tradeoffs. Shared rooms may provide social interaction, companionship, and cognitive stimulation—potentially valuable given high rates of loneliness and social isolation among nursing home patients (Trybusińska and Saracen, 2019; Zhang et al., 2023). However, roommates may also impose costs: sleep disruption, noise, exposure to behavioural disturbances (particularly common among patients with AD/ADRD), loss of privacy, and elevated risk of infectious-disease transmission (Brown et al., 2021; Konetzka, Grabowski, and Mor, 2024).

These tradeoffs have motivated ongoing policy debates about room configuration standards. The Biden Administration's 2022 nursing home reform agenda explicitly proposed accelerating the phase-out of rooms with three or more patients and promoting single-occupancy rooms (The White House, 2022). States such as Massachusetts, Michigan, and Ohio have introduced incentives to encourage the adoption of private rooms. Germany has taken an even more ambitious approach: several German federal states have implemented regulatory mandates requiring nursing homes to convert multi-patient rooms into predominantly single-occupancy rooms over a defined transition period (Herr, Lückemann, and Reichert, 2025). However, these policies remain controversial. Single rooms substantially increase construction costs and may reduce social interaction, with uncertain net effects on patient well-being.

Resolving this debate requires understanding not just whether room sharing affects outcomes on average, but also whether effects depend on roommate characteristics. This heterogeneity is particularly relevant for patients with AD/ADRD. Dementia-related symptoms such as confusion, agitation, wandering, and repetitive vocalizations can be distressing to cognitively intact roommates. At the same time, social interaction and peer monitoring may benefit individuals with AD/ADRD themselves (Nichols, 2014). Specialized dementia care units—which often feature single rooms or modified environments—are designed to address these needs but remain uncommon: only 4.5% of nursing homes beds are in special care units (Joyce et al., 2018). Most patients with AD/ADRD thus reside in general care units, where room assignments determine their daily social environment.

2.1.2 Room Assignment as a Policy Lever

Room assignment policies represent a potentially important but understudied margin for improving nursing home outcomes. Unlike many proposed interventions, optimizing room assignments requires primarily better information and revised protocols rather than additional staffing or costly facility modifications. If peer effects vary systematically with roommate characteristics, facilities can improve outcomes by better matching patients—for instance, avoiding particularly poor matches, or prioritizing single rooms for patients most likely to be harmed by room sharing.

Despite the potential policy relevance, empirical evidence on peer effects in nursing homes is extremely limited. Prior research has examined facility-level quality measures (Grabowski, Gruber, and Angelelli, 2008), the effectiveness of specialized dementia units (Joyce et al., 2018), and infection transmission in shared rooms (Brown et al., 2021; Konetzka, Grabowski, and Mor, 2024). However, we are not aware of prior work that credibly identifies causal effects of individual roommate characteristics on patient outcomes.

2.2 Data

To quantify the causal effect of the assignment of nursing home rooms on patient mortality, we combine three primary data sources. First, we use the Minimum Data Set (MDS) to measure room assignment as well as patient and peer health characteristics at admission. A key feature of our identification strategy is that we focus on assignments to three room types: private rooms, shared rooms where the roommate has Alzheimer's disease or related dementias (AD/ADRD), and shared rooms where the roommate does not have AD/ADRD. Second, we link the data to Medicare claims data to measure patient mortality, our primary outcome of interest. Third, we incorporate facility-level data from the On-Line Survey, Certification, and Reporting (OSCAR) system to validate our room measures and explore potential mechanisms.

2.2.1 Minimum Data Set (MDS)

The MDS provides standardized patient assessments for all patients in Medicare or Medicaid-certified nursing homes, regardless of payer source. Mandated by federal law, these assessments are conducted at admission, discharge, and quarterly intervals during nursing home stays.[2] The data have been widely used in prior research and offer detailed information on patient demographics, health status, and functional limitations. We use version 2.0, which was in effect during our 2000-2010 study period.

[2] In addition, during our study period, assessments were also required for Medicare-covered patients at days 14, 30, and 60.

Room Assignment Measures

The key innovation in our empirical strategy is the use of a room identifier variable contained in MDS assessments. By combining room identifiers with admission and discharge dates, we can reconstruct the daily composition of patients in each nursing home room throughout our sample period.

Constructing room-level variables. We combine the room identifier with admission and discharge dates to build a nursing- home-by-room-by-day database, which tracks the daily room composition at all nursing home rooms. This allows us to identify the type of room each patient is assigned to at admission, as well as the room composition at the nursing home she is assigned to. Finally, we construct room capacity as the maximum number of patients co-residing in the room at any point during that year. This allows us to measure available beds as the difference between room capacity and current utilization.

Validation of room measures. We conduct two validation exercises to ensure our room identifier captures meaningful variation in actual room assignments.

First, Appendix Figure A.1 compares the number of distinct room identifiers in the MDS to bed counts from OSCAR facility data. Consistent with expectations, the number of rooms is smaller than the number of beds and increases approximately linearly with facility size, with a slope consistent with most rooms containing one or two beds.

Second, Appendix Table **??** documents substantial homophily in roommate pairings. Patients are far more likely to share rooms with others of the same gender, similar age, and similar health status—patterns consistent with room assignments reflecting deliberate matching decisions rather than random measurement error.

Taken together, this evidence suggests that our room measures capture meaningful information on actual room assignments. While measurement error cannot be entirely ruled out, we note that such error would likely attenuate our estimates toward zero, making our findings conservative lower bounds on the true peer effects.

Defining Treatment: Room Types. Our identification strategy exploits quasi-random variation in assignment to three room types, defined by the presence and health status of roommates at the time of admission. We classify each admission into one of three mutually exclusive categories:

1. **Empty room**: The patient is assigned to a room without roommates. This includes private and empty multi-bed rooms.

2. **Shared room with AD/ADRD roommate**: The patient is assigned to share a room with at least one roommate who has a documented AD/ADRD diagnosis.

3. **Shared room without AD/ADRD roommate**: The patient is assigned to share a room with at least one roommate, none of whom have documented AD/ADRD.

These room types represent the ”treatment” in our empirical analysis. In rooms with multiple roommates, we classify the room as having an AD/ADRD roommate if *any* roommate has the diagnosis. We measure room composition at the time of the focal patient's admission, which is pre-determined and not affected by the health of the newly admitted patient.

Health and Demographic Measures

We construct measures of both focal patients' own health characteristics and their roommates' characteristics at the time of admission.

Dementia diagnosis. Our primary measure of peer cognitive impairment is an indicator for whether a roommate has a diagnosis of Alzheimer's disease or related dementias (AD/ADRD). The MDS dementia diagnosis fields capture physician-documented conditions that are clinically active and relevant to the patient's current care plan. These fields reflect facility-recognized diagnoses rather than staff observations or cognitive test scores alone. Prior validation studies show that they provide highly specific markers of dementia but may miss milder or undocumented cases (Niznik et al., 2025). We construct an analogous measure to characterize focal patients' own dementia status.

Cognitive, functional status, and patient demographics. We measure cognitive impairment using the Cognitive Performance Scale (CPS), a validated summary measure ranging from 0 (intact) to 6 (very severe impairment) and activities of Daily Living (ADL) where higher values indicate greater difficulty with tasks like eating, dressing, and toileting. These measure have been used as measures of cognitive and physical impairment in numerous studies (Grabowski, Gruber, and Angelelli, 2008; Rahman, Norton, and Grabowski, 2016; Cornell et al., 2019). The MDS also provides information on patient demographics (age, sex, race), and prior living situation (community, hospital, another nursing home). Together with the measures of cognitive and functional status, we use these variables as controls in some specifications and to test for effect heterogeneity.

2.2.2 Medicare Claims Data

We link MDS records to Medicare beneficiary files and claims data using unique beneficiary identifiers. This linkage allows us to track health outcomes even after patients are discharged from nursing homes, avoiding potential selection bias from conditioning on continued nursing home residence.

Our primary outcome is mortality within 90 days of nursing home admission, measured using the death date recorded in the Medicare Beneficiary Summary File. We focus on 90-day mortality because it is a well-measured, policy-relevant outcome that has been used extensively in health economics research (Chandra et al., 2016b; Finkelstein, Gentzkow, and Williams, 2021). In robustness checks, we examine alternative time horizons (30-day, 180-day, and 360-day mortality). We also extract 27 chronic condition indicators from the Medicare Chronic Conditions Warehouse, including conditions such as diabetes, heart failure, and chronic obstructive pulmonary disease. We use these measures to control for baseline health differences in robustness exercises.

2.2.3 OSCAR Facility Data

We merge OSCAR data to obtain annual facility-level characteristics for each nursing home in our sample. OSCAR collects these data through periodic surveys and certification inspections conducted by state agencies. We use OSCAR data for two purposes. First, as described above, we use bed counts to validate our room identifier measures (Appendix Figure A.1). Second, we use facility characteristics—including total beds, ownership type (for-profit, nonprofit, government), staffing ratios, and the presence of specialized Alzheimer's or dementia care units—to explore potential mechanisms underlying our main results.

2.2.4 Sample Construction

We construct our analysis sample using MDS 2.0 data from 2000–2010 for twelve large states: California, New York, Florida, Texas, Pennsylvania, Ohio, Illinois, New Jersey, Massachusetts, Indiana, Michigan, and North Carolina. These states were selected based on data availability and quality, and collectively account for approximately 60% of U.S. nursing home patients.[3]

By linking the MDS to Medicare claims data, we restrict the sample to newly admitted patients aged 65 and older enrolled in traditional (fee-for-service) Medicare during our study period. This restriction ensures complete ascertainment of mortality outcomes while covering the majority of nursing home admissions. During our sample period, about 80% of the elderly were enrolled in traditional Medicare plans (Gold et al., 2011). Importantly, while we restrict focal patients to Medicare beneficiaries, we observe the full set of roommates regardless of payer source.

We further limit the data to each patient's first observed nursing home stay. Room assignment in subsequent stays may follow different rules (e.g., due to bed-hold policies)[4], and later stays are observed only for patients who survive the initial admission, potentially biasing the sample. Our final dataset contains 2.6 million nursing home stays across more than 7,200 facilities and 480,000 unique rooms. Further details on the sample construction are provided in Appendix Section B.

2.3 Summary Statistics

The columns 1–4 of Table 1 provide summary statistics on baseline characteristics for the full sample, patients assigned to rooms without a roommate, patients assigned roommate(s) with AD/ADRD, and patients assigned roommate(s) none of whom have AD/ADRD respectively. In order to focus on within-nursing-home-year differences in baseline characteristics between patients assigned to different types of rooms, these variables are residualized of nursing-home-by-year fixed effects (with the overall mean added back). Columns

[3]Calculations are based on patient counts in 2024 from Nursing Home Compare available at https://www.kff.org/state-category/providers-service-use/, last accessed November 18th, 2025.

[4]If a patient is discharged to a hospital and expected to return, the nursing home may hold her bed during the interim.

5 and 6 show difference in means for the baseline characteristic for different subsamples along with standard errors for these differences clustered at the nursing home level.

The first six rows show summary statistics for some basic demographic variables. We observe that more than half of patients are female, the average age is 78, and that the majority are white, have less than a Bachelor's degree, are post-acute care, and are not on Medicaid at admission.

The last four rows show baseline health of patients. For easier interpretation, we start by combining the many health variables into a single summary index: baseline mortality risk. We do so by considering the regression of 90-day mortality on the demographic variables, each of the 27 chronic conditions (as defined by the CMS Chronic Conditions Data Warehouse), fixed effects for CPS and ADLs, indicators for whether the patient is assigned to an empty room and whether the patient is assigned a roommate with AD/ADRD, as well as nursing-home-by-year fixed effects. We define baseline mortality risk as the predicted values from this regression, net of the room effects and the nursing-home-by-year fixed effects, and we obtain these predicted values using k-fold cross-validation with $k = 10$.

In columns 5 and 6, we observe that patients assigned to rooms without a roommate as well as patients assigned to a room where at least one roommate has AD/ADRD tend to have a slightly higher baseline mortality risk than patients assigned to rooms with roommates who do not have AD/ADRD. We also see statistically significant differences in CPS scores, ADL scores, and number of chronic conditions for patients assigned to different types of rooms, suggesting that room assignment is not random and that naive OLS regressions will not recover the causal effect of roommates on future health.

3 Empirical Strategy

Our primary goal is to quantify the causal effects of nursing home room assignments on patients' health outcomes. Our main specification is:

$$Y_{irjt} = \beta_0 + \beta_1 \, \text{Empty}_{r(i),j,t(i)} + \beta_2 \, \text{AD}_{r(i),j,t(i)} + \delta_{j,y(t)} + \varepsilon_{irjt}, \tag{1}$$

where Y_{irjt} is a future health outcome (90-day mortality in our main specifications) for patient i admitted to room r in nursing home j on day t, $\text{Empty}_{r(i),j,t(i)}$ indicates whether i

Table 1: Summary Statistics

	Full Sample	Assigned No Roommate	Assigned Roommate(s) with AD/ADRD	Assigned Roommate(s) All Without AD/ADRD	Diff: (2)-(4)	Diff: (3)-(4)
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.625	0.608	0.644	0.617	-0.012***	0.031***
	(0.470)	(0.476)	(0.463)	(0.472)	(0.001)	(0.001)
Education: Bachelor's Degree	0.083	0.089	0.079	0.083	0.008***	-0.004***
	(0.260)	(0.277)	(0.250)	(0.260)	(0.001)	(0.000)
Age	78.243	78.007	78.692	77.937	0.033	0.844***
	(11.589)	(11.741)	(11.251)	(11.813)	(0.024)	(0.025)
Race: White	0.812	0.817	0.809	0.813	0.005***	-0.004***
	(0.310)	(0.301)	(0.310)	(0.315)	(0.001)	(0.001)
Post-Acute Care	0.656	0.652	0.643	0.670	-0.018***	-0.029***
	(0.406)	(0.409)	(0.416)	(0.393)	(0.001)	(0.001)
Baseline Mortality Risk	0.154	0.156	0.155	0.153	0.003***	0.002***
	(0.102)	(0.104)	(0.101)	(0.103)	(0.000)	(0.000)
AD/ADRD	0.299	0.291	0.340	0.264	0.028***	0.085***
	(0.437)	(0.432)	(0.462)	(0.412)	(0.001)	(0.001)
Cognitive Performance Scale	1.802	1.781	1.924	1.695	0.085***	0.248***
	(1.550)	(1.541)	(1.571)	(1.525)	(0.004)	(0.004)
Physical Impairment (ADLs)	14.704	14.637	14.721	14.725	-0.068***	0.038***
	(7.131)	(7.164)	(7.186)	(7.060)	(0.015)	(0.015)
Number of Chronic Conditions	6.563	6.528	6.589	6.557	-0.042***	0.036***
	(3.862)	(3.840)	(3.845)	(3.889)	(0.008)	(0.007)
Number of Observations	2,517,826	551,527	968,835	997,464	–	–

Notes: Columns 1–4 report means (with standard deviations in parentheses) for: (1) all patients; (2) patients assigned to empty rooms; (3) patients assigned to a room with at least one roommate with AD/ADRD; and (4) patients assigned to a non-empty room where no roommate has AD/ADRD. All variables are residualized on nursing-home-by-year fixed effects with the grand mean added back. Baseline mortality risk is estimated using 10-fold cross-validation from regressions of 90-day mortality on baseline demographics and health covariates, controlling for nursing-home-year fixed effects and room assignment. Columns 5 and 6 report differences in means relative to column 4, with standard errors clustered at the nursing home level.

is assigned to an empty room, and $\text{AD}_{r(i),j,t(i)}$ indicates whether at least one of i's roommates at admission has AD/ADRD. The excluded group comprises patients assigned to a shared room (in which none of the roommates has a diagnosis of AD/ADRD), and $\delta_{j,y(t)}$ are nursing-home-by-year fixed effects.

The coefficients of interest are β_1 and β_2. We interpret them as intent-to-treat estimates based on initial assignment. Since room transitions over the course of a nursing home stay are likely endogenous, we abstract from them in this paper.

Identification Challenges: There are at least three empirical challenges that may bias OLS estimates of β_1 and β_2. First, the reflection problem (Manski, 1993). In peer effects models with contemporaneous outcomes, it is difficult to separate whether patient A affects patient B or vice versa. We address this concern by measuring roommate characteristics (AD/ADRD status) at the time of patient i's admission, ensuring these measures are predetermined and cannot be influenced by i's subsequent health trajectory.

Second, endogenous selection out of nursing homes. In principle, the length of nursing home stay is endogenous to patient and provider incentives and to changes in health (Hackmann, Pohl, and Ziebarth, 2024; Einav, Finkelstein, and Mahoney, 2025). This introduces econometric challenges when the measurement of outcomes is contingent on residing in a nursing home, e.g. outcomes collected in the MDS. To avoid that, we only use MDS data collected at patient i's admission and instead focus on 90-day mortality recorded in Medicare claims data, which we can measure regardless of nursing home attendance.

Third, endogenous room assignment within facilities. While nursing-home year fixed effects control for quality differences between nursing homes as well as compositional differences in patients attending a specific nursing home, a valid concern is that assignment to rooms within nursing homes is not random. For instance, the evidence presented in Table 1 suggests that patients with cognitive impairments are more likely to be assigned a room shared with a patient who has a diagnosis of AD/ADRD. To address this concern, we use an instrumental variable strategy which leverages high-frequency (daily) variation in room composition.

3.1 Endogenous room assignments

We instrument room assignment – $\text{Empty}_{r(i),j,t(i)}$ and $\text{AD}_{r(i),j,t(i)}$ – with the share of rooms with a vacant bed at the nursing home that are empty – $s^{\text{Empty}}_{j,t(i)}$ – and the share of rooms with a vacant bed that have an existing patient with AD/ADRD – $s^{\text{AD}}_{j,t(i)}$ – at the time of i's admission. Intuitively, if all available rooms are empty (or already have an existing patient with AD/ADRD) on the day of admission, then any newly admitted patient will be assigned to an empty room (respectively, a roommate with AD/ADRD). Precisely, we measure $s^{\text{Empty}}_{j,t(i)}$ and $s^{\text{AD}}_{j,t(i)}$ on the day prior to i's admission. This ensures that our instrumental variables are not mechanically affected by i's admission. Hence, the first stage equations are:

$$\text{Empty}_{r(i),j,t} = \phi_{j,y(t)} + \tau_1 \, s^{\text{Empty}}_{j,t(i)} + \tau_2 \, s^{\text{AD}}_{j,t(i)} + \eta_{irjt} \tag{2a}$$

$$\text{AD}_{r(i),j,t} = \xi_{j,y(t)} + \alpha_1 \, s^{\text{Empty}}_{j,t(i)} + \alpha_2 \, s^{\text{AD}}_{j,t(i)} + \nu_{irjt} \tag{2b}$$

Following equation (1), we again include nursing home-by-year fixed effects ($\phi_{j,y(t)}, \xi_{j,y(t)}$).

To provide context, Appendix Figure A.2 presents histograms of our instruments. We observe wide variation in the overall distribution of these instruments in panels (a) and (b) of the figure, although the share of available rooms that are empty is typically smaller than the share of available rooms with a patient that has AD/ADRD. However, this variation may also be due to differences between nursing-home-years (e.g., the share of rooms with a patient who has AD/ADRD may be higher in nursing homes' that have Alzheimer's units), and our IV strategy leverages within-nursing-home-year variation. Indeed, panels (c) and (d) show that the within-nursing-home-year distributions of the instruments are tighter[5]. We observe that substantial variation in the instruments remain, due to volatility in admissions and discharges (both in volume and composition).

[5]Specifically, we plot the distributions of the variables residualized of nursing-home-by-year fixed effects with the overall mean added back. Note that it is possible for these residualized variables to have support outside of $[0, 1]$: for example, if the average share of available rooms with a patient with AD/ADRD is 0.5 in the full sample and 0.75 in a nursing home in a given year, then if this share drops to zero for that nursing home in a given day during that year, the value of the normalized variable would be $0 - 0.75 + 0.5 = -0.25$. We drop the very small number of cases where the residualized variables lie outside of $[0, 1]$ from the histograms for easier visualization.

3.1.1 Identification

To interpret the 2SLS estimates of β_1 and β_2 as the causal effects of being assigned an empty room or a room shared with a patient with an AD/ADRD diagnosis, our instrumental variables need to be relevant and conditionally independent of potential outcomes. Furthermore, in the presence of heterogeneous treatment effects, in order to interpret β_1 and β_2 as properly weighted averages of treatment effects for different types of patients, the relationship between treatments and instruments needs to exhibit average conditional monotonicity and no cross-effects (Bhuller and Sigstad, 2024). In this section, we summarize empirical evidence supporting the validity of these identifying assumptions.

Relevance: Our approach requires that our instruments have strong predictive power for the endogenous variables.[6] We assess this by examining the relationships between $\text{Empty}_{r(i),j,t}$ and $s^{\text{Empty}}_{j,t(i)}$ as well as between $\text{AD}_{r(i),j,t}$ and $s^{\text{AD}}_{j,t(i)}$, conditional on nursing-home-by-year fixed effects.

The blue dots in Figure 1 present the first stage results. Figure 1a presents results from the first stage regression in a binscatter plot for one of our instruments: the share of available rooms occupied with at least one patient with AD/ADRD at the time (day) of admission, $s^{AD}_{j,t(i)}$ (holding the other instrument fixed). The blue line (circle markers) shows a positive linear relationship between the probability of being assigned to a roommate with AD/ADRD (vertical axis) and the instrument $s^{AD}_{j,t(i)}$ (horizontal axis), $\alpha_2 > 0$, providing strong support for instrumental relevance. Moreover, the relationship is remarkably linear, which supports the linear specification in equation (2b). We find that a 1 percentage point increase in the share of available rooms with a patient diagnosed with AD/ADRD increases the probability of being assigned a patient diagnosed with AD/ADRD by 1.16 percentage points, with a t-statistic of more than 200.

The blue dots in Figure 1b present corresponding evidence for the second instrument: the share of available empty rooms, $s^{Empty}_{j,t(i)}$ (holding $s^{AD}_{j,t(i)}$ fixed). The blue line (circle markers) shows again a positive linear relationship between the probability of being assigned to an empty room (vertical axis) and $s^{Empty}_{j,t(i)}$ (horizontal axis), $\alpha_1 > 0$, providing again strong support for instrumental relevance. The relationship is again remarkably linear, support-

[6] Specifically, in the case with two endogenous variables and two excluded instruments, the 2×2 matrix of first stage coefficients must have rank 2.

ing the linear specification in equation (2a). We find that a 1 percentage point increase in the share of available empty rooms increases the probability of being assigned an empty room by 0.97, with a t-statistic greater than 170.

Conditional Independence: For our instrument to be valid, our instruments must be uncorrelated with patients' potential outcomes, conditional on the baseline controls including nursing-home-by-year fixed effects. While the institutional features described above support conditional independence, this assumption may still be violated if patients delay their admission until a more preferred room type becomes available, or if they change their nursing home choice because their preferred room type is unavailable.

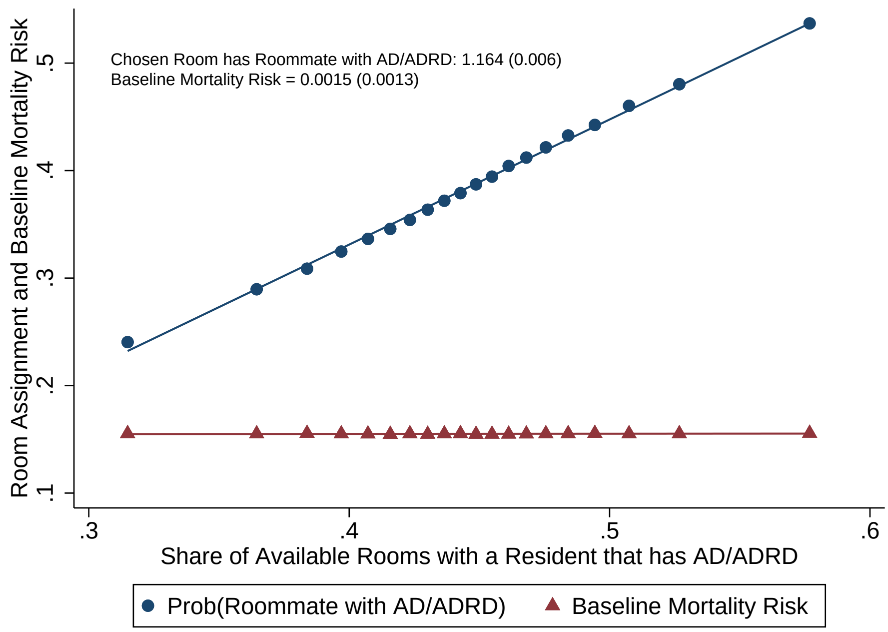
To address this concern, we present two sets of empirical evidence which provide strong support for the conditional independence assumption. First, we show that patients' observed characteristics are well balanced across the instruments, conditional on our baseline controls. As shown by the red dots and lines in Figure 1, predicted mortality constructed on baseline patient characteristics at admission is remarkably stable across variation in the instruments. Specifically, the red line in Figure 1a plots the relationship between observable baseline mortality risk of patient i (vertical axis) and the share of empty rooms occupied with a patient diagnosed with AD/ADRD. The coefficient estimate is small in magnitude and not statistically significant at the five percent significance level.

The red line in Figure 1b revisits this for our second instrument, the share of empty rooms, $s_{j,t}^{Empty}$. Again, the coefficient estimate is small, and even though the estimate is statistically significant at the five percent level, given our later results that being assigned to an empty room reduces 90-day mortality on average, if anything this small bias from imbalance in baseline mortality risk may lead us to *underestimate* the effect of being assigned to an empty room.

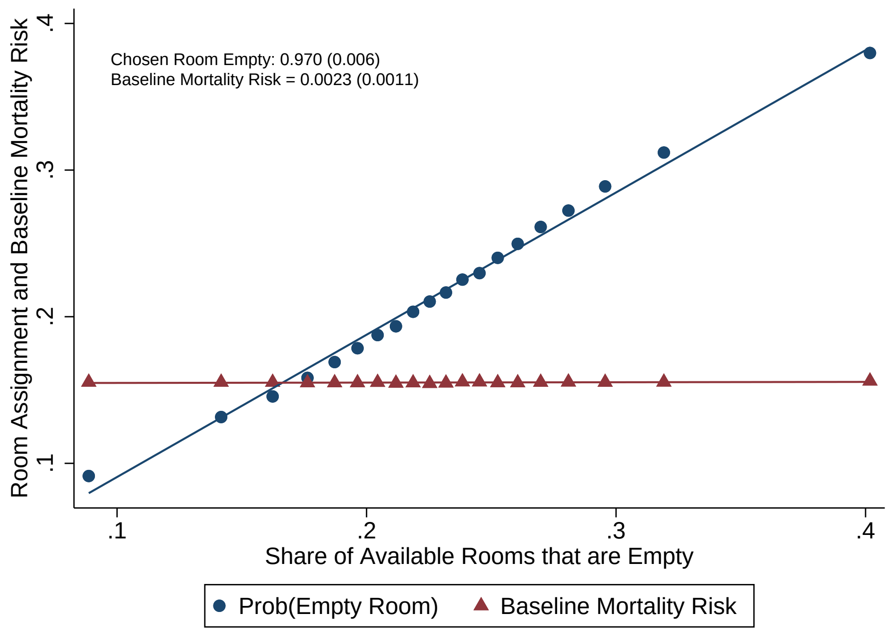
To probe the conditional independence assumption for our instruments further, Appendix Figure A.3 explores balance for 40 baseline characteristics with respect to our instrument, comparing the magnitude of potential imbalances with the reduced form estimates for the instruments. Specifically, we regress 90-day mortality as well as each of the baseline characteristics on each instrument and nursing-home-by-year fixed effects (without controlling for the other instrument so that the reduced form coefficients are easier to

Figure 1: First stage and Balance in Mortality Risk

(a) Assignment to Roommate with AD/ADRD Diagnosis



(b) Assignment to Room Without a Roommate



Notes: Panel (a) plots the probability of being assigned a roommate with AD/ADRD and baseline mortality risk as a function of the share of available rooms with a patient that has AD/ADRD at the time of admission, controlling for nursing-home-by-year fixed effects and the share of available rooms that are empty at the time of admission. Panel (b) plots the probability of being assigned to an empty room as a function of the share of available rooms that are empty at the time of admission, controlling for nursing-home-by-year fixed effects and the share of available rooms that have a patient with AD/ADRD at the time of admission. Coefficient estimates and standard errors are based on linear regressions with the same outcomes and regressors, with standard errors clustered at the nursing home level. Each circle and triangle corresponds to a bin containing 5 percent of the sample (20 bins in total).

interpret).[7] To ensure comparability of the coefficients' scale and sign across regressions for different variables, we multiply the coefficient estimate (and standard error) by the dependent variable's association with 90-day mortality.

The results in Appendix Figure A.3 shows that for both instruments, the magnitudes of the reduced form estimates (shown at the top of the two panels) are far larger than the coefficient estimates for baseline characteristics. The coefficient estimates for a number of baseline characteristics are statistically different from zero at the five percent significance level, but tend to have inconsistent signs. Indeed, when we combine the baseline characteristics into baseline mortality risk, this risk measure is essentially conditionally uncorrelated with $s^{AD}_{j,t(i)}$ in panel (a), and although its correlation with $s^{empty}_{j,t(i)}$ is barely statistically significant at the five percent significance level in panel(b), the sign of this correlation is the opposite of the reduced form estimate, suggesting that even if this imbalance results in bias, it is likely to attenuate the IV estimate of β_1 towards zero.

Average Conditional Monotonicity and No Cross-Effects: In the presence of heterogeneous treatment effects, we need additional assumptions to interpret the IV estimates as properly weighted average of treatment effects across different patients (i.e., with non-negative weights and zero weights on the "wrong" treatment effect).[8] Bhuller and Sigstad (2024) show that in contrast to IV with a single endogenous variable where only monotonicity is required, in IVs with multiple endogenous variables two assumptions are required: average conditional monotonicity and no cross-effects.

To describe these two assumptions, note that we can generate rotation-invariant instruments such that each instrument is associated with a single treatment using the predicted values from regressions of each treatment variable on both instruments (and the controls). We denote these instruments by P^{Empty} and P^{AD}. Average conditional monotonicity re-

[7] In particular, the IV estimate with a single endogenous variable and a single instrument is given by the ratio of the reduced form coefficient to the first stage coefficient, whereas in the case with two endogenous variables and two instruments, the IV estimates are given by the inverse of the 2×2 matrix of first stage coefficients multiplied by the reduced form coefficients, which is harder to show graphically.

[8] Specifically, one can show that the IV estimate for β_1 can be expressed as:

$$\beta_1 = \mathbb{E}[\omega_1^S \beta_1^S + \omega_2^S \beta_2^S],$$

where S indexes response types and $\beta_k^s \equiv \mathbb{E}[\beta_k | S = s]$. Bhuller and Sigstad (2024) show that under the conditional independence and first stage assumptions, the weights on β_1^s sum to one while the weights on β_2^s sum to zero. However, additional assumptions are required so that weights are proper, i.e., that the weights on β_1^s are non-negative, and the weights on β_2^s are zero. A similar expression applies for β_2.

quires that P^{Empty} (respectively, P^{AD}) only affects Empty (AD) weakly positively, while no cross-effects requires that P^{Empty} does not affect AD (conditional on P^{AD}) and vice versa.

These assumptions have a testable implication similar to that for monotonicity in IV with one treatment variable: if we divide the data into different subsamples based on baseline characteristics, the regressions of Empty (AD) on P^{Empty} and P^{AD} (and controls) should result in a positive coefficient estimate on P^{Empty} (P^{AD}) and a zero estimate (up to statistical uncertainty) on P^{AD} (P^{Empty}). Appendix Figure A.4 shows histograms of coefficient estimates with Empty and AD as the dependent variables in panels (a) and (b) respectively, for over 3,300 subsamples defined based on 42 baseline characteristics and their pairwise interactions. The blue bars (respectively, red bars) correspond to estimates for the transformed instrument that matches (does not match) the dependent variable. We observe that all the blue bars are large and positive, supporting average conditional monotonicity, whereas all of the red bars are small and close to zero, supporting the no cross-effects assumption.

4 Results

4.1 Average Effects

Table 2 shows IV and OLS estimates of equation (1) for the full sample. Column 1 shows that being assigned to an empty room at admission reduces 90-day mortality by 1.3 percentage points relative to being assigned a roommate, an 8.4 percent reduction relative to the mean. Column 2 shows that this effect depends on the characteristics of roommates patients are assigned to: relative to being assigned roommates that do not have dementia (the excluded group), being assigned to an empty room at admission reduces 90-day mortality by 0.7 percentage points while being assigned a roommate that has dementia increases mortality by 1.4 percentage points. In both IV specifications, the first-stage F-statistics are in the tens of thousands, so weak instruments are unlikely to be an issue. Combined, this implies that moving from a room shared with a patient with AD/ADRD to an empty room reduces 90-day mortality by 2.1 percentage points. To put this estimate into perspective, we note that Cheng (2023) estimates nursing home value added to 90-day survival and finds a standard deviation of 2 percentage points. That means that moving from a room shared with a patient with AD/ADRD to an empty room reduces 90-day mortality is simi-

Table 2: Average Effects of Room Assignment on 90-Day Mortality

	IV		**OLS**	
	(1)	(2)	(3)	(4)
Assigned to Room with No Roommate	-0.013***	-0.007**	0.005***	0.006***
	(0.003)	(0.003)	(0.001)	(0.001)
Assigned to Roommate with AD/ADRD		0.014***		0.001
		(0.003)		(0.001)
F-statistic	36,383	25,835	–	–
Dependent Variable Mean	0.155	0.155	0.155	0.155
Number of Observations	2,612,318	2,612,318	2,612,318	2,612,318

Notes: This table shows IV and OLS estimates of equation (1). Controls include nursing home-by-year fixed effects. Standard errors are clustered at the nursing home level.

lar to falling by one standard deviation in the value added distribution of nursing homes. Put differently, this suggests that there are not only large quality differences between nursing homes Olenski and Sacher (2024); Einav, Finkelstein, and Mahoney (2025); Hackmann, Rojas, and Ziebarth (2025) but also within nursing homes.

Columns 3 and 4 of Table 2 show OLS estimates for comparison. The estimate in column 3 suggests that being assigned to an empty room *increases* 90-day mortality, while column 4 suggests that relative to being assigned roommates who do not have AD/ADRD, being assigned to an empty room *increases* mortality while being assigned a roommate with AD/ADRD has no statistically significant effect on mortality. The sign of the empty-room estimates contrasts with the IV results, indicating that room assignment is correlated with unobserved health determinants and underscoring the need for our IV strategy to identify causal effects. Moreover, the small OLS estimates for sharing a room with a patient with AD/ADRD may reflect attenuation from measurement error in room assignment—another concern mitigated by the IV approach.

Post-acute care versus long-stay patients: Our large sample encompasses nursing home stays for all traditional Medicare beneficiaries, including stays originating from a hospital. These patients primarily require post-acute care and typically have relatively short stays. The sample also includes long-stay patients, typically admitted from community-based care settings, who have long-term care needs and typically longer stays. A natural question is whether the effect sizes differ between these populations.

The first two columns in panel A of Table 3 splits the sample into these populations. Interestingly, in column 1 we find no evidence for reductions in mortality from assignment to an empty room for post-acute care patients (relative to assignment to roommates without AD/ADRD). This is consistent with the results in (McWilliam, 2025), who also find no improvement in this population. One reason could be that the value of an empty room is smaller for shorter stays. However, our results also show that this depends on the point of comparison. We find again a large positive effect on patient mortality when assigned to a room shared with a patient diagnosed with AD/ADRD. This means that an empty room assignment does result in lower 90-day mortality when compared against assignment to a room shared with a patient with AD/ADRD.

Column 2 repeats the exercise for long-stay patients. Here, we find significant reductions in 90-day mortality from assignment to an empty room. The effect size of assignment to a room shared with a patient with AD/ADRD is similar to the effect for post-acute care patients.

4.2 Substitutability and Complementarities in Health Care Production

To explore peer-patient health interactions, Panel A of Table 3 examines patients with and without AD/ADRD at admission (columns 3 and 4). The results reveal asymmetric effects. Empty room assignment reduces mortality for patients without AD/ADRD but increases mortality for those with AD/ADRD. Similarly, having a roommate with AD/ADRD increases mortality for patients who also have AD/ADRD, but has no effect on cognitively intact patients. These patterns imply that cognitively intact roommates are particularly beneficial for patients with AD/ADRD—reducing their 90-day mortality by 5.1 percentage points—while patients without AD/ADRD show no mortality benefit from cognitively intact roommates relative to private rooms. These findings suggest potential gains from pairing patients with AD/ADRD with cognitively intact roommates, as we discuss below.

One explanation is that cognitively healthy roommates provide informal monitoring—alerting staff when help is needed—which is especially valuable for patients with AD/ADRD who often lack situational awareness. Without a cognitively intact roommate present, this protective monitoring is lost. Conversely, roommates who themselves have AD/ADRD cannot reliably provide this monitoring function, explaining why patients with dementia benefit specifically from cognitively healthy roommates rather than from any roommate. This monitoring mechanism aligns with findings in Nichols (2014), who argues that patients with significant cognitive impairment often cannot recognize when they need assistance or reliably summon help. Private-room designs eliminate the roommate who might alert staff when a patient is in distress—a loss that is particularly consequential for individuals with dementia who depend on others to recognize their needs.

An alternative explanation could be that patients with AD/ADRD demand particular attention, crowding out staff support for their roommates. While this can explain the negative effect of an AD/ADRD patient on a peer AD/ADRD patient it is different to reconcile with the absence of a negative effect on peers without AD/ADRD.

Independent of the underlying mechanism, the presented findings suggest that the room composition has large implications for average mortality. We illustrate this in three stylized examples in Appendix Figure C.5 that vary the health composition of patients and the availability of room types. Our findings suggest that average nursing home mortality varies by up to 2.8 percentage points under different room assignments.

4.3 Robustness Checks

We test the robustness of our main results in several ways. First, we test whether our IV estimates remain stable as we sequentially include additional control variables. Specifically, we flexibly control for occupancy, to account for potential changes in patient composition at higher or lower occupancy (Gandhi, 2023; Hackmann, Pohl, and Ziebarth, 2024), county-by-year-by-month fixed effects to account for potential market level trends in patient arrivals or exits, day of the week fixed effects to capture differences in admissions and discharges on weekdays or weekends, room fixed effects to isolate differences in patient composition from general room amenities, and rich baseline health and demographic characteristics. The results presented in Appendix Table A.2 show that the IV estimates remain stable as we sequentially add different sets of controls.

To better understand sources of bias in the OLS estimates, in Appendix Table A.3 we conduct the same exercise using the OLS estimates, sequentially adding different sets of controls. Interestingly, while the sign of the OLS estimates differ from the IV estimates in the baseline specification, in the most saturated specifications, the OLS estimates flip sign and aligns with the IV estimates, suggesting that some of the differences can be explained for by selection on observables.

Second, we test whether the estimated treatment effects are specific to the time horizon for mortality that we chose. Appendix Figure A.4 shows our main IV estimates with 30-day, 90-day, 180-day and 360-day mortality as the outcome. We observe that treatment effects for 30-day mortality have the same sign as our main estimates based on 90-day mortality, but are typically smaller in magnitude. This is consistent with peer effects among roommates growing over time. In addition, we also observe that the estimated effects on mortality tend to persist until (at least 360 days). This suggests that our peer effects estimates are not driven by short-run displacement or harvesting, whereby roommate exposure merely shifts the timing of deaths by a few weeks or months. Rather, the effects seem to reflect persistent changes in patients' health trajectories.

Third, we derive bounds for our IV estimates allowing for potential violations of conditional independence under various assumptions about the unobservables, similar in spirit to the methods proposed in Altonji, Elder, and Taber (2005), Conley, Hansen, and Rossi (2012), and Oster (2019). At a high level, given assumptions on the maximum amount of variation in the outcome that the instruments, controls and unobservables can explain (R^2_{max}) and the relative degree of selection on observables and unobservables (δ) in the reduced form regression, bounds for the reduced form coefficients are identified, which can be translated to bounds for the treatment effects by inverting the matrix of first stage coefficients. Under Oster's recommended choices of R^2_{max} and δ (which are based on the R^2 from the reduced form estimates and the movement in reduced form coefficients as we include additional controls), the bounds we obtain for the effects of being assigned to an empty room or a roommate with AD/ADRD on mortality are [-0.0104, -0.0027] and [0.0121, 0.0153] respectively. Appendix Section E explains these bounds in greater detail, we show additional bounds in Appendix Table E.5, as well as values of (R^2_{max}, δ) required for violations of IV assumptions to completely explain our non-zero treatment effect estimates in Appendix Figure E.7.

Fourth, we address the issue of "power asymmetry" in 2SLS estimation when t-tests are used inference, as pointed out by Keane and Neale (2023).[9] The IV estimates in Table 2 indicate that on average, being assigned to an empty room reduces 90-day mortality while being assigned a roommate with AD/ADRD increases mortality. On the other hand, the OLS estimates biases us against finding these effects. Hence, if anything, the power asymmetry in t-tests for inference in 2SLS is likely to bias us *against* finding significant IV estimates. Moreover, simulations in Keane and Neale show that the power asymmetry is ameliorated as the strength of the first stage increases, with power being roughly symmetric when the first stage F-statistic is around 105. In our setting, the first stage F-statistic is orders of magnitude larger, so we expect power asymmetry to be even less of an issue.

4.4 Room Assignments and Average Mortality

These findings raise a natural policy question: which room assignment rules would minimize overall mortality? In this subsection, we derive intuitive optimality conditions and use them to quantify the potential mortality reductions from optimal room assignments under various scenarios.

Formally, we consider a room-specific health production function

$$U = f(i, i', r) \tag{2}$$

that depends on patient health i and peer health i' at admission. r denotes the room. Assignment to an empty room is captured by $i' = 0$. We assume that total health production is additive in room specific average mortality and define $\mathcal{A}$ to be the set of all feasible allocations, where each allocation $A \in \mathcal{A}$ consists of a set of such tuples satisfying capacity and feasibility constraints (e.g., each patient is assigned to exactly one room, room capaci-

[9] While Keane and Neale (2023) study 2SLS with a single treatment variable, we assume that a similar intuition holds for the case with two treatment variables, one with each treatment variable.

ties, etc.). Defining U as a positive outcome, here 90-day survival, we solve the following problem:

$$A^* \in \arg\max_{A \in \mathcal{A}} \sum_{(i,i',r)\in A} f(i,i',r). \tag{3}$$

Appendix Section D considers a simple version of this model with two types of patients (with or without AD/ADRD), two types of rooms (private room with one bed or shared room with two beds), and where the number of patients is equal to the nursing home's bed capacity, N. The feasible set of allocations is then characterized by the share of patients with AD/ADRD, π, and the share rooms that are private, p: $\mathcal{A}(\pi, p, N)$. The allocation of patients to rooms is then fully characterized by two decision variables: x^0_{AD} and s_{AN}, which are the number of patients with AD/ADRD in private rooms and the number of shared rooms that are mixed (one patient with AD/ADRD, 'A', and one without, 'N'), respectively.

The solution to this model boils down to a linear programming problem, which after simplifying accounting balances, can be written as:

$$x^{0,*}_{\text{AD}}, s^*_{\text{AN}} = \underset{x^0_{\text{AD}}, s_{\text{AN}} \in \mathcal{A}(\pi,p,C)}{\arg\max} \quad \psi x^0_{\text{AD}} + \theta s_{\text{AN}}, \tag{4}$$

where:

$$\psi \equiv (u_{\text{AD},0} - u_{\text{AD,AD}}) - (u_{\text{No},0} - u_{\text{No,No}}),$$

$$\theta \equiv (u_{\text{AD,No}} + u_{\text{No,AD}}) - (u_{\text{AD,AD}} + u_{\text{No,No}}),$$

where, $u_{i,i'}$ denotes 90-day survival for patient i with AD/ADRD (AD) or without (No), when assigned to an empty room (0), or a room shared with or without a roommate with AD/ADRD.

The key parameters that determine optimal allocation are ψ and θ, which both have intuitive economic interpretations: ψ is the difference in privacy premium between the two types of patients. The estimates from Table imply (after flipping signs to capture effects on survival) $\psi = (-1.4 + 5.1) - (1.4 - 0) = 2.3$ percentage points. θ compares the gains from having mixed rooms relative to segregated rooms. Rearranging terms, our estimates suggest $\theta = (u_{AD,No} - u_{AD,AD}) - (u_{No,AD} - u_{No,No}) = 5.1 + 0.6 = 5.7$ percentage points.

When θ is positive, the production function is submodular and roommates' health are substitutes, so we typically want to maximize the number of shared rooms that are mixed.[10] On the other hand, when θ is negative, the production function is supermodular and roommates' health are complements, in which case it is optimal to have segregated rooms, and private rooms are allocated to the type with the greater privacy premium (which depends on the sign of ψ).

We illustrate the difference between the survival-maximizing assignment rule and alternative assignment rules that impose random assignment or segregation by health in Appendix Figure D.6, for different room configurations (specifically, varying the fraction of rooms that are private, p) and different patient compositions (varying the fraction of the patient population that has dementia, π). For all values of p and π considered, average mortality is lowest under the survival-maximizing assignment rule, followed by the random assignment rule, and highest under segregated rooms. Differences between assignment rules in mortality rates tend to be largest when the patient population is relatively mixed ($\pi \approx 0.5$) and most rooms are shared ($p \approx 0$): in this case, 90-day mortality rate is about 1.5 (3) percentage points higher under random assignment (respectively, segregated rooms) compared to the survival-maximizing assignment rule.

The empirically relevant case is visualized Figure D.6a. We estimate that $\pi = 30\%$ of patients are diagnosed with AD/ADRD in our sample, and that approximately $p = 25\%$ of rooms are private. At baseline, we estimate an average 90-day mortality rate of 14.9%[11]. Under the optimal assignment rule, this could be reduced to 14.1%, implying a 0.8 percentage point reduction in average mortality (5.4% reduction relative to baseline).

4.5 Mechanisms

In this section, we consider several mechanisms that may help reconcile the large peer effects discussed above.

[10] The exception to this rule is when the privacy premium outweighs the gains from mixing ($|\psi| \geq \theta$), in which case we may want to sacrifice some mixed rooms in order to allocate more patients of the type with a higher privacy premium to private rooms.

[11] This differs from the average mortality rate of 15.5% reported in Table 2 because we are dropping patients with missing or invalid values of AD/ADRD at admission.

Vaccination: While the results in Panel A of Table 3 suggest that roommates without AD/ADRD may play a support role which can be especially valuable for patients with AD/ADRD, they do not explain why being assigned to an empty room leads to lower mortality rates for patients without AD/ADRD. Panel B explores one possible explanation: lower risk of disease contagion.

To explore this possibility, we estimate the effect of being assigned to a room without a roommate for patients who did or did not receive flu vaccinations in columns 1 and 2, and PPV vaccinations in columns 3 and 4. While the effects of being assigned to an empty room are similar for patients who did or did not receive flu vaccines, the effect is smaller for patients who received a PPV vaccine compared to patients who did not receive a PPV vaccine. This provides suggestive evidence that disease contagion may partially explain the reduction in 90-day mortality for patients assigned to empty rooms.

Staff Shortages: Panel C explores whether peer effects vary with nurse staffing levels, an important determinant of patient health (Lin, 2014; Harrington et al., 2020). Nurse shortages can compromise monitoring capacity, resulting in patient neglect and higher mortality (Friedrich and Hackmann, 2021). If so, patients without AD/ADRD may provide crucial support in understaffed facilities. We estimate room effects for nursing homes with above- and below-median staffing (columns 1-2), where staffing equals the sum of z-scores for RN, LPN, and CNA levels. Effects are concentrated in facilities with below-median staffing levels, suggesting that healthy peers and adequate staffing can act as substitutes in health production.

Alzheimer's units: Finally, in columns 3 and 4 of panel C, we test whether peer effects vary by the presence of an Alzheimer's unit. Nursing homes with specialized Alzheimer's units are better equipped to support AD/ADRD patients through enhanced staff training, specialized floor plans designed for monitoring and safety, and tailored environmental features (Joyce et al., 2018), which may dampen positive spillovers from cognitively intact roommates. Consistent with this, the final two columns show that the effect of being assigned a roommate with AD/ADRD (versus without) is small and insignificant in facilities with Alzheimer's units, but remains large and highly significant in facilities without such units, which comprise the vast majority of observations. This pattern suggests that healthy peers and specialized capital infrastructure can act as substitutes in supporting patients with cognitive impairments.

5 Conclusion

This paper provides causal evidence that peer effects among nursing home roommates generate mortality impacts comparable in magnitude to facility quality differences. Assignment to a roommate with AD/ADRD, relative to a private room, increases 90-day mortality by 2.1 percentage points (14% of baseline)—equivalent to receiving care at a nursing home one standard deviation worse in quality.

Critically, we find substantial heterogeneity in these peer effects. Patients with AD/ADRD benefit from cognitively healthy roommates but not from private rooms, suggesting that peers provide important monitoring and support. In contrast, patients without AD/ADRD are unaffected by roommate cognitive health but benefit from privacy. These patterns are most pronounced in facilities with below-median staffing and without specialized dementia units, pointing to substitutability between peer health and institutional resources.

These findings suggest a novel, low-cost channel for improving patient outcomes: strategic room assignment. A simple assignment rule that places AD/ADRD patients with cognitively healthy roommates while prioritizing private rooms for cognitively healthy patients could reduce overall mortality by 0.8 percentage points—without additional staff or facility modifications. This represents a 5.2% reduction in 90-day mortality rates. While strategic peer assignment has been explored in education settings (Carrell, Sacerdote, and West, 2013), it has received little attention in healthcare, where research has focused primarily on patient-provider matching.

Several caveats merit attention. First, our estimates reflect effects under current assignment practices; optimal assignment may generate different effect sizes through equilibrium effects on peer composition. Second, we focus on mortality; effects on other outcomes (e.g., cognitive and physical health, quality of life, family satisfaction, and staff burden) warrant investigation, although rigorous statistical analysis of room effects for other patient outcomes needs to contend with competing risks (given the effects we find for mortality). Third, we focus on the effects of being assigned roommates with or without AD/ADRD given the increasing prevalence and costs of these diseases; in reality, peer effects along

Table 3: Heterogeneity in Room Effects

Panel A: Post-Acute Care and Cognitive Health

	Post-Acute	Not Post-Acute	AD/ADRD	No AD/ADRD
	(1)	(2)	(3)	(4)
Assigned to Room with No Roommate	-0.002	-0.011**	0.013**	-0.014***
	(0.004)	(0.005)	(0.006)	(0.004)
Assigned to Roommate with AD/ADRD	0.012***	0.011**	0.051***	-0.006
	(0.004)	(0.006)	(0.006)	(0.004)
F-statistic	16523	11072	10034	17734
Dependent Variable Mean	0.157	0.151	0.161	0.144
Number of Observations	1,710,649	897,040	750,156	1,764,818

Panel B: Vaccination Status

	Flu Vaccine	No Flu Vaccine	PPV Vaccine	No PPV Vaccine
	(1)	(2)	(3)	(4)
Assigned to Room with No Roommate	-0.014**	-0.015***	-0.011**	-0.021***
	(0.006)	(0.006)	(0.005)	(0.005)
First Stage F-statistic	9454	10622	14617	10595
Dependent Variable Mean	0.115	0.125	0.120	0.126
Number of Observations	594,933	707,962	769,303	635,708

Panel C: Facility Characteristics

	Above-Median Staffing	Below-Median Staffing	Alzheimer's Unit	No Alzheimer's Unit
	(1)	(2)	(3)	(4)
Assigned to Room with No Roommate	-0.004	-0.010**	-0.001	-0.008**
	(0.004)	(0.005)	(0.007)	(0.003)
Assigned to Roommate with AD/ADRD	0.011**	0.017***	-0.001	0.017***
	(0.004)	(0.005)	(0.007)	(0.004)
First Stage F-statistic	14390	13402	5239	21040
Dependent Variable Mean	0.154	0.156	0.152	0.156
Number of Observations	1,343,186	1,269,056	562,962	2,049,280

Notes: Panel A splits the sample between patient groups. Columns 1 and 2 consider patients admitted from a post-acute care hospital or not. Columns 3 and 4 consider patients with and without a diagnosis of AD/ADRD at admission, dropping patients with missing or invalid values for initial AD/ADRD status. Panel B splits the sample by patient vaccination status. Panel C splits the sample into nursing home with above or below-median staffing levels and with or without an Alzheimer's special care unit.

other dimensions may exist. Fourth, the simple assignment policy considered in this paper is static (assigning all patients simultaneously), whereas in practice patients arrive and exit stochastically.

Future work will extend this analysis by developing a dynamic structural model to characterize optimal assignment policies accounting for the full distribution of patient characteristics, stochastic arrival patterns, and capacity constraints. Additionally, examining mechanisms more directly—through data on falls, medication adherence, and staff response times—would strengthen causal interpretation and guide implementation.

More broadly, our findings highlight that peer environments constitute a potentially important but understudied determinant of health outcomes in institutional settings. As policymakers seek to improve care quality in nursing homes, hospitals, and other congregate care facilities, strategic peer assignment may offer meaningful gains at minimal cost.

References

Abdulkadiroğlu, Atila, Joshua Angrist, and Parag Pathak. 2014. "The elite illusion: Achievement effects at Boston and New York exam schools." *Econometrica* 82 (1):137–196.

Alsan, Marcella, Owen Garrick, and Grant Graziani. 2019. "Does Diversity Matter for Health? Experimental Evidence from Oakland." *American Economic Review* 109 (12):4071–4111.

Alsan, Marcella and Marianne Wanamaker. 2018. "Tuskegee and the health of black men." *The quarterly journal of economics* 133 (1):407–455.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools." *Journal of Political Economy* 113 (1):151–184.

Bhuller, Manudeep and Henrik Sigstad. 2024. "2SLS with Multiple Treatments." *Journal of Econometrics* 242 (1):105785.

Brown, Kevin A, Aaron Jones, Nick Daneman, Adrienne K Chan, Kevin L Schwartz, Gary E Garber, Andrew P Costa, and Nathan M Stall. 2021. "Association between nursing home crowding and COVID-19 infection and mortality in Ontario, Canada." *JAMA internal medicine* 181 (2):229–236.

Carrell, Scott E., Richard L. Fullerton, and James E. West. 2009. "Does Your Cohort Matter? Measuring Peer Effects in College Achievement." *Journal of Labor Economics* 27 (3):439–464.

Carrell, Scott E., Bruce Sacerdote, and James E. West. 2013. "From natural variation to optimal policy? The importance of endogenous peer group formation." *Econometrica* 81 (3):855–882.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson. 2016a. "Health care exceptionalism? Performance and allocation in the US health care sector." *American Economic Review* 106 (8):2110–2144.

———. 2016b. "Productivity dispersion in medicine and manufacturing." *American Economic Review* 106 (5):99–103.

Cheng, Alden. 2023. *Demand for quality in the presence of information frictions: Evidence from the nursing home market*. SSRN.

Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi. 2012. "Plausibly exogenous." *Review of Economics and Statistics* 94 (1):260–272.

Cornell, Paul Y. et al. 2019. "Do report cards predict future quality? The case of skilled nursing facilities." *Journal of Health Economics* 66:208–221.

Doyle, Joseph, John Graves, and Jonathan Gruber. 2019. "Evaluating measures of hospital quality: Evidence from ambulance referral patterns." *Review of Economics and Statistics* 101 (5):841–852.

Einav, Liran, Amy Finkelstein, and Neale Mahoney. 2025. "Producing Health: Measuring Value Added of Nursing Homes." *Econometrica* 93 (4):1225–1264.

Einav, Liran, Amy Finkelstein, Neale Mahoney, and James C Okun. 2025. "Racial Differences in Nursing Home Value Added." Tech. rep., National Bureau of Economic Research.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2021. "Place-based drivers of mortality: Evidence from migration." *American Economic Review* 111 (8):2697–2735.

Friedrich, Benjamin U and Martin B Hackmann. 2021. "The returns to nursing: Evidence from a parental-leave program." *The Review of Economic Studies* 88 (5):2308–2343.

Gandhi, Ashvin. 2023. "Picking Your Patients: Selective Admissions in the Nursing Home Industry." Working paper, SSRN.

Gandhi, Ashvin, Andrew Olenski, Krista Ruffini, and Karen Shen. 2024. "Alleviating Worker Shortages Through Targeted Subsidies: Evidence from Incentive Payments in Healthcare." *Review of Economics and Statistics* .

Geweke, John, Gautam Gowrisankaran, and Robert J Town. 2003. "Bayesian inference for hospital quality in a selection model." *Econometrica* 71 (4):1215–1238.

Gold, Marsha, Gretchen Jacobson, Anthony Damico, and Tricia Neuman. 2011. "Medicare Advantage enrollment market update." Tech. rep., Mathematica Policy Research. URL https://www.kff.org/wp-content/uploads/2013/01/8228.pdf.

Grabowski, David C., Jonathan Gruber, and Joseph J. Angelelli. 2008. "Nursing home quality as a common good." *The Review of Economics and Statistics* 90 (4):754–764.

Greenwood, Brad N, Rachel R Hardeman, Laura Huang, and Aaron Sojourner. 2020. "Physician–patient racial concordance and disparities in birthing mortality for newborns." *Proceedings of the National Academy of Sciences* 117 (35):21194–21200.

Hackmann, Martin B. 2019. "Incentivizing Better Quality of Care: The Role of Medicaid and Competition in the Nursing Home Industry." *American Economic Review* 109 (5):1684–1716.

Hackmann, Martin B, R Vincent Pohl, and Nicolas R Ziebarth. 2024. "Patient versus provider incentives in long-term care." *American Economic Journal: Applied Economics* 16 (3):178–218.

Hackmann, Martin B, Juan S Rojas, and Nicolas R Ziebarth. 2025. "Creative Financing and Public Moral Hazard: Evidence from Medicaid and the Nursing Home Industry." Tech. rep., National Bureau of Economic Research.

Harrington, Charlene, Mary Ellen Dellefield, Elizabeth Halifax, Mary Louise Fleming, and Debra Bakerjian. 2020. "Appropriate nurse staffing levels for US nursing homes." *Health services insights* 13:1178632920934785.

Herr, Annika, Markus Lückemann, and Arne R. Reichert. 2025. "The rise of person-centered care: Effects of single-room nursing home quotas on long-term care." Working Paper 734, Hannover Economic Papers.

Hull, Peter. 2018. "Estimating hospital quality with quasi-experimental data." *Available at SSRN 3118358* .

Hurd, Michael D., Pierre-Carl Michaud, and Susann Rohwedder. 2017. "Distribution of lifetime nursing home use and of out-of-pocket spending." *Proceedings of the National Academy of Sciences* 114 (37):9838–9842.

Joyce, N. R., T. G. McGuire, S. J. Bartels, S. L. Mitchell, and D. C. Grabowski. 2018. "The impact of dementia special care units on quality of care: An instrumental variables analysis." *Health Services Research* 53 (5):3657–3679.

Kaiser Family Foundation. 2025. "5 Key Facts About Nursing Facilities and Medicaid." URL https://www.kff.org/medicaid/5-key-facts-about-nursing-facilities-and-medicaid/. https://www.kff.org/medicaid/5-key-facts-about-nursing-facilities-and-medicaid/.

Keane, Michael and Timothy Neale. 2023. "Instrument strength in IV estimation and inference: A guide to theory and practice." *Journal of Econometrics* 235 (2):1625–1653.

Konetzka, R Tamara, David C Grabowski, and Vincent Mor. 2024. "Four Years And More Than 200,000 Deaths Later: Lessons Learned From The COVID-19 Pandemic In US Nursing Homes: Study examines COVID-19 and nursing home deaths." *Health Affairs* 43 (7):985–993.

Lin, Haizhen. 2014. "Revisiting the relationship between nurse staffing and quality of care in nursing homes: An instrumental variables approach." *Journal of health economics* 37:13–24.

Manski, Charles F. 1993. "Identification of endogenous social effects: The reflection problem." *The Review of Economic Studies* 60 (3):531–542.

Matthews, Kevin A, Wei Xu, Anne H Gaglioti, James B Holt, Janet B Croft, Dominic Mack, and Lisa C McGuire. 2019. "Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015–2060) in adults aged 65 years." *Alzheimer's & Dementia* 15 (1):17–24.

McWilliam, Dianne. 2025. "The Effect of Private vs Shared Rooms on Nursing Home Resident Health Outcomes." *Available at SSRN 5654692* .

Mukamel, Dana B. et al. 2023. "Dementia care is widespread in U.S. nursing homes; facilities with the most dementia patients may offer better care." *Health Affairs* 42 (6):795–803.

Nichols, Jeffrey. 2014. "Private rooms not always a better place for residents." *Caring for the Ages* 15 (2):3.

Niznik, Joshua D, Florentia E Sileanu, Xinhua Zhao, Kelvin Tran, Laura C Hanson, Alan Kinlaw, Thomas R Radomski, Alexa Ehlert, Sydney Springer, Binxin Cao et al. 2025. "A Comparison of Measures for Identifying Possible Dementia in Veterans Affairs Nursing Home Residents." *Journal of the American Medical Directors Association* 26 (4):105481.

Olenski, Andrew and Szymon Sacher. 2024. "Estimating nursing home quality with selection." *Review of Economics and Statistics* :1–31.

Oster, Emily. 2019. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of Business Economic Statistics* 37 (2):187–204.

Rahman, Momotazur, Edward C. Norton, and David C. Grabowski. 2016. "Do hospital-owned skilled nursing facilities provide better post-acute care quality?" *Journal of Health Economics* 50:36–46.

Sacerdote, Bruce. 2001. "Peer effects with random assignment: Results for Dartmouth roommates." *The Quarterly Journal of Economics* 116 (2):681–704.

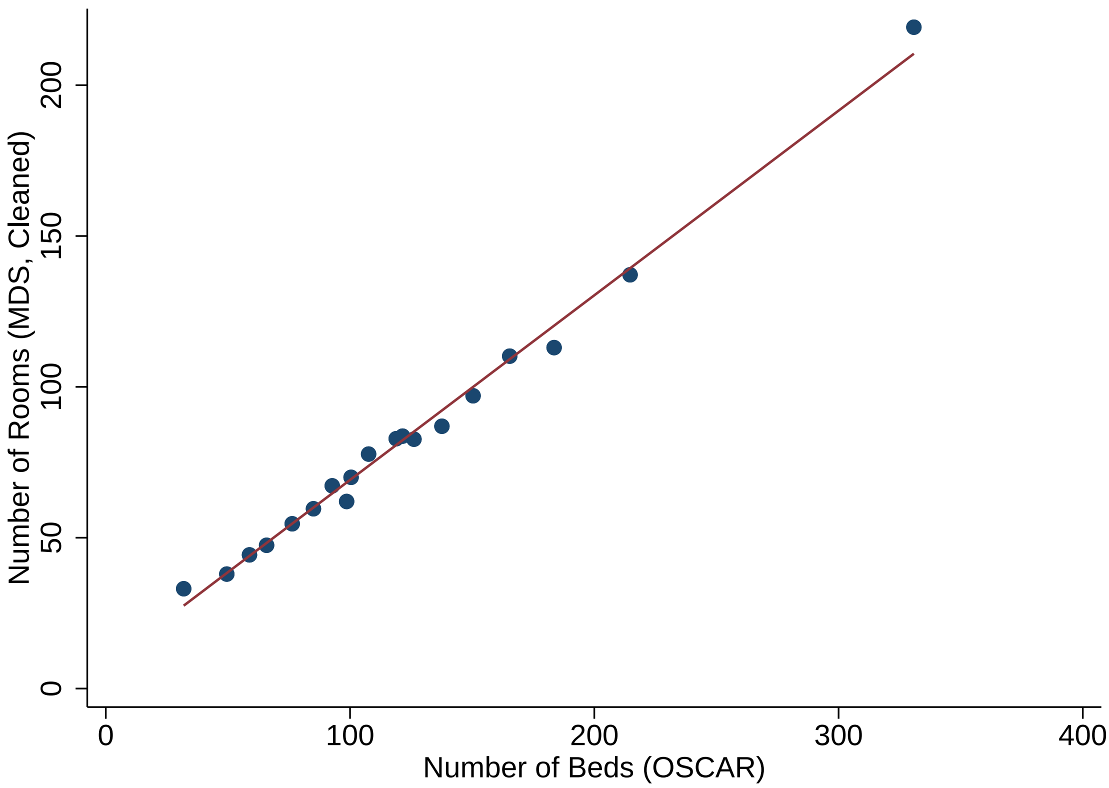
The White House. 2022. "Fact sheet: Protecting seniors by improving safety and quality of care in the nation's nursing homes." URL https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2022/02/28/fact-sheet-protecting-seniors-and-people-with-disabilities-by-improving-safety-and-quality-of-care-in-the-nations-nursing-homes/.

Trybusińska, Dorota and Agnieszka Saracen. 2019. "Loneliness in the context of quality of life of nursing home residents." *Open medicine* 14 (1):354–361.

Zhang, Deping, Qizhen Lu, Li Li, Xiaofeng Wang, Hongqi Yan, and Zijian Sun. 2023. "Loneliness in nursing homes: A qualitative meta-synthesis of older people's experiences." *Journal of Clinical Nursing* 32 (19-20):7062–7075.

A Appendix Figures and Tables

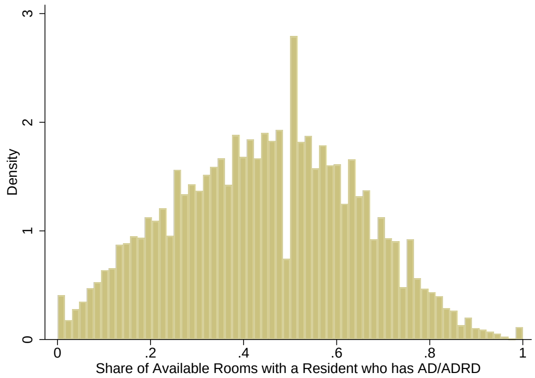
Figure A.1: Relationship Between Number of Rooms and Number of Beds



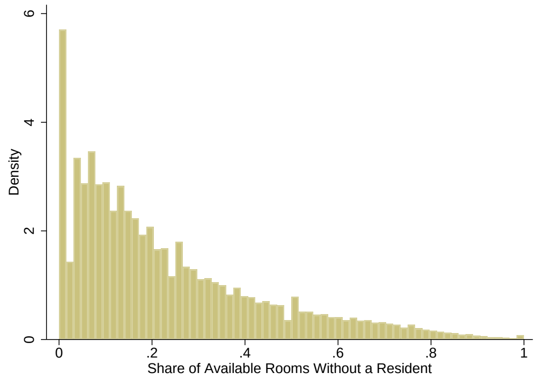
Notes: This figure plots the relationship between number of rooms in each nursing home each year based on the room identifier in the MDS data, and the number of beds in the same nursing-home-year according to the OSCAR data.

Figure A.2: Distribution of Share of Available Rooms with a Patient that has AD/ADRD or Without a Patient

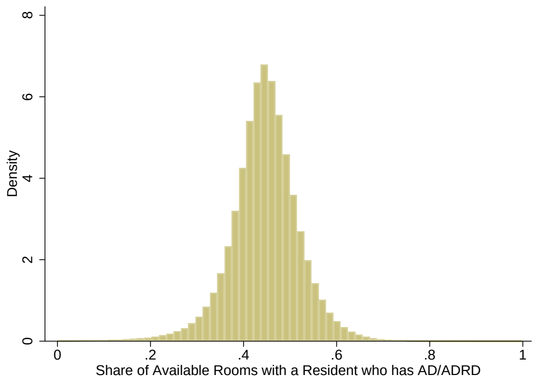
(a) Rooms with a Patient that has AD/ADRD



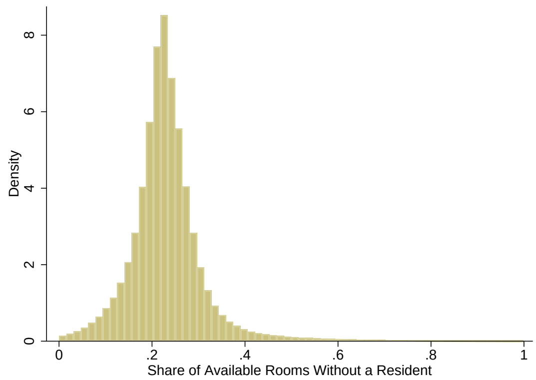
(b) Rooms Without a Patient



(c) Rooms with a Patient that has AD/ADRD (Within-Nursing-Home-Year)



(d) Rooms Without a Patient (Within-Nursing-Home-Year)

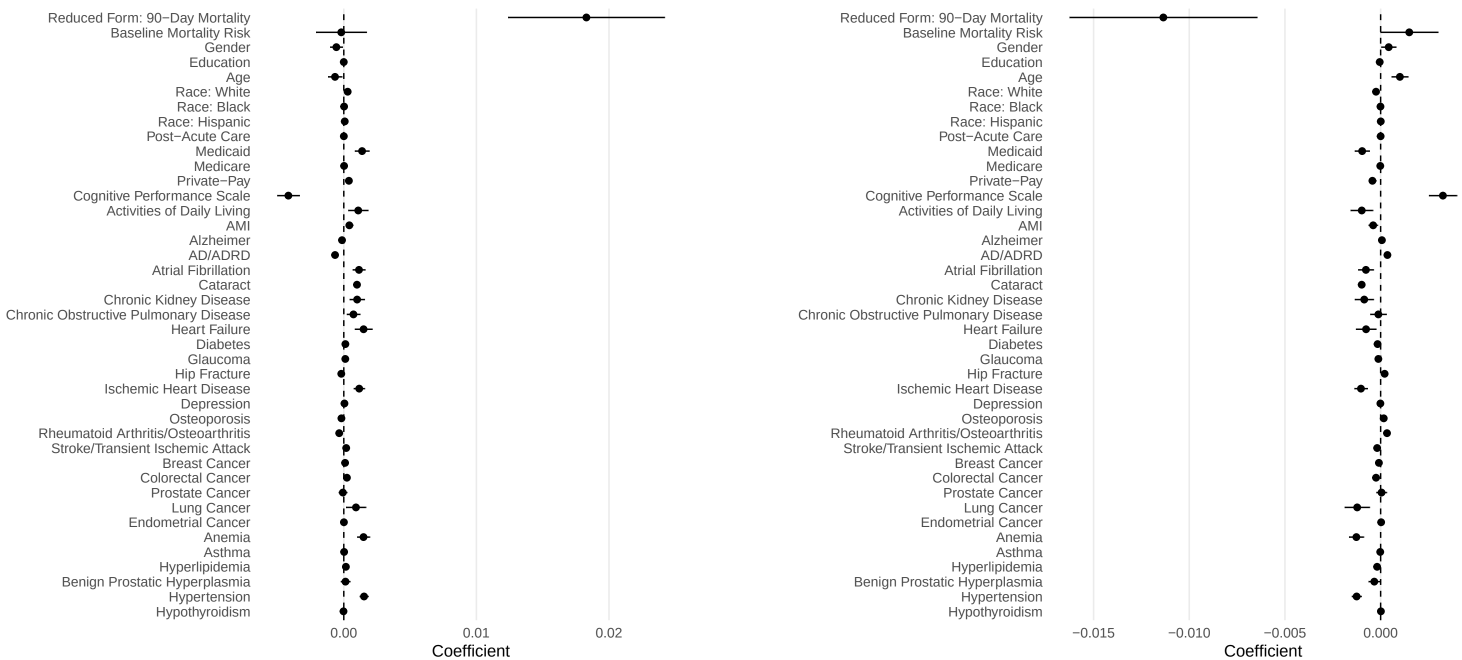


Notes: This figure plots histograms showing the distributions of our instruments – share of available rooms with a patient that has AD/ADRD and share of available rooms that are empty – both unconditionally in panels (a) and (b), and conditional on nursing-home-by-year fixed effects in panels (c) and (d). Specifically, for the within-nursing-home-year distributions of our instruments, we first residualize the instruments of nursing-home-by-year fixed effects, before adding back their overall means.

Figure A.3: Reduced Form and Conditional Independence

(a) Share of Available Rooms with a Patient that has AD/ADRD

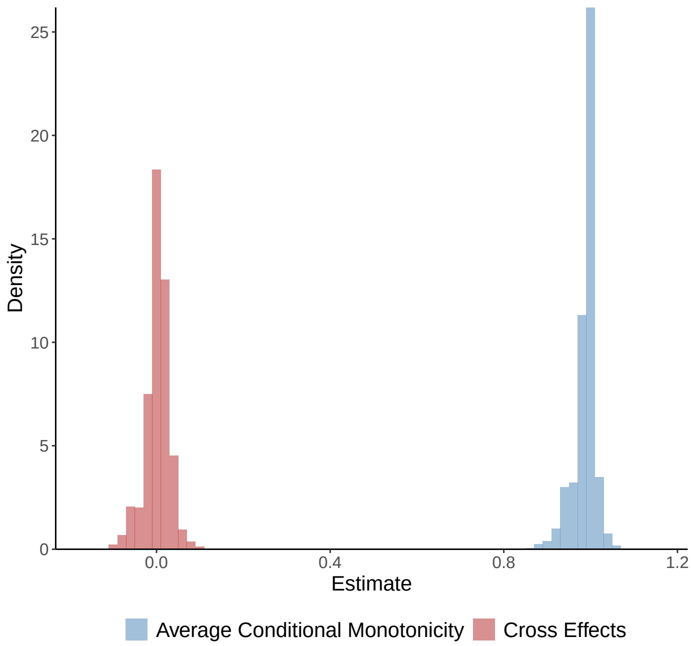
(b) Share of Available Rooms Without a Roommate



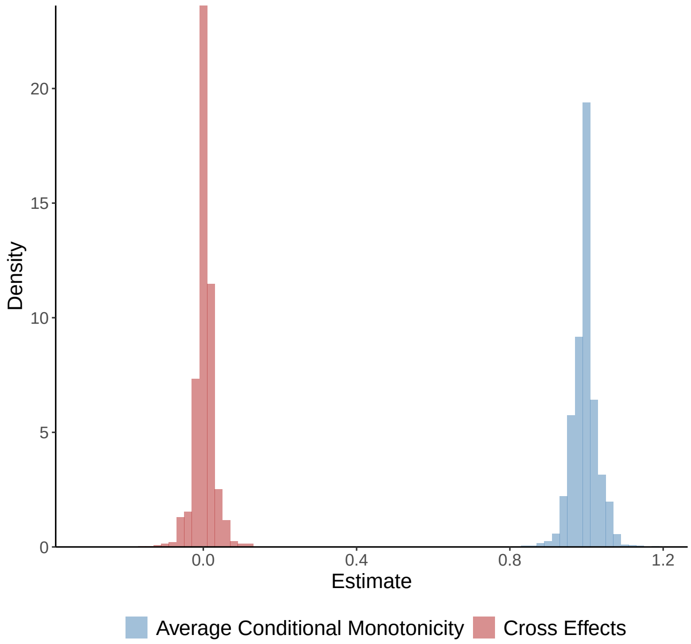
Notes: This figure provides evidence on the conditional independence assumption by comparing regression estimates of baseline characteristics on the instruments (from the second row onwards) with the reduced form estimate (of the outcome on the instrument) in the first row. Specifically, panel (a) and (b) plot coefficient estimates and 95 percent confidence intervals from regressions of 90-day mortality (in the first row) and baseline patient characteristics (in the remaining rows) on the share of available rooms with a patient that has AD/ADRD and the share of available rooms that are empty respectively, controlling for nursing-home-by-year fixed effects, and with standard errors clustered at the nursing home level.

Figure A.4: Average Conditional Monotonicity and No Cross-Effects

(a) Empty as the Dependent Variable



(b) AD as the Dependent Variable



Notes: These figure shows histograms of coefficient estimates from regressions of Empty and Empty on P^{Empty} and P^{AD} (and controls) for different subsamples. The dependent variable in panel (a) is Empty and the dependent variable in panel (b) is AD. The subsamples are defined based on 42 baseline characteristics as well as their pairwise interactions. We consider all possible pairwise combinations except for when the sample restriction leads to less than 10,000 observations, resulting in a total of more than 3,300 different subsamples. The blue bars (respectively, red bars) correspond to estimates on the transformed instrument that matches (does not match) the dependent variable.

Table A.1: Relationship Between Resident and Roommate Characteristics at Admission

	AD/ADRD (1)	Female (2)	Post-Acute Care (3)	Medicaid (4)	Bachelor's Degree (5)	Age $\leq$ 80 (6)	Race: White (7)
Share of Roommates with AD/ADRD	0.094*** (0.002)						
Share of Roommates that are Female		0.642*** (0.003)					
Share of Roommates that are Post-Acute Care			0.051*** (0.001)				
Share of Roommates on Medicaid				0.056*** (0.001)			
Share of Roommates with Bachelor's Degree					0.014*** (0.002)		
Share of Roommates Aged $\leq$ 80						0.048*** (0.001)	
Share of Roommates that are White							0.071*** (0.002)
Nursing Home $\times$ Year Fixed Effects	X	X	X	X	X	X	X
Number of Observations	2,087,100	2,086,611	2,087,100	2,087,100	2,087,100	2,087,100	2,087,100
R-squared	0.093	0.397	0.280	0.219	0.099	0.115	0.380

Notes: This table shows regressions of patient characteristics at admission on the average value of these characteristics among the patient's roommate(s) at the time of the patient's admission, and nursing-home-by-year fixed effects. The sample is limited to patients who are assigned a roommate, and standard errors are clustered at the nursing home level.

B Details on Sample Construction

Cleaning the room identifier. The room identifier requires substantial cleaning because its format varies across nursing homes and within nursing homes over time. In cleaning the room identifier, We address the following main issues.

First, the variable sometimes records bed assignments rather than room assignments. For example, entries like ”15-1” or ”15/A” likely indicate bed 1 or bed A within room 15. In such cases, we strip the bed suffix to obtain the room number.

Second, some entries are too coarse or invalid to be useful. For instance, entries like ”SOUTH” (referring to an entire wing) or ”ROOM” (a placeholder) do not identify specific rooms. We drop patients with such room identifiers from the analysis.

Third, the way a nursing home refers to the same physical room may change over time due to facility expansions, reorganizations, or inconsistencies in data entry. A room initially labeled ”15” might later be recorded as ”A15.” To avoid treating these as distinct rooms—which would create ”phantom” rooms that appear available but do not actually exist—we define the set of available rooms separately for each nursing home and year.[12]

Finally, we drop rooms with maximum recorded occupants exceeding four patients, as these likely reflect coding errors or institutional settings (e.g., hospital wards) that differ meaningfully from typical nursing home rooms. Our main results are robust to alternative occupancy thresholds.[13]

Missing Discharge Dates. Some patient stays are missing discharge assessments (which nursing homes are required to fill in when patients leave the facility, be it due to discharge to hospital/community or due to death). In particular, no further assessments are recorded for a patient for a given stay, with the last (non-discharge) assessment being at least 150 days before the end of the sample. This threshold is chosen in accordance with the definition adopted by the CMS for ”active” patients.[14] In such cases, we assume that the patient was discharged one quarter after the last available assessment (given that nursing homes are required to fill in assessments at least quarterly), or the day before the beginning of the patient's next stay (at the same nursing home or another nursing home) if this occurs less than one quarter after the last assessment for the previous stay.

Our final analysis sample contains 2.6 million nursing home admissions across 7,200 facilities and 480,000 unique rooms, with 551,527 (22%) assigned to private rooms, 968,835 (38%) to shared rooms with AD/ADRD roommates, and 997,464 (40%) to shared rooms without AD/ADRD roommates.

C Room Assignments and Aggregate Mortality: A few examples

Figure C.5 illustrates how room assignment affects aggregate mortality through several stylized examples. Example A considers three patients—two diagnosed with AD/ADRD (red) and one without (green)—and two rooms. In case A1, the patient without AD/ADRD occupies the empty room; in A2, a patient with AD/ADRD does. Switching from A1 to A2 reduces average 90-day mortality by 2.7 percentage points. While the patient without

[12] Specifically, we first construct daily room occupancies for the entire sample period. Then, we define a room as available at a point in time t if the room's maximum occupancy in that year ($y(t)$) is at least one, and the occupancy at time t is less than the maximum occupancy observed in $y(t)$.

[13] Results available from authors upon request.

[14] https://data.cms.gov/resources/facility-level-minimum-data-set-frequency-methodology. Accessed November 22, 2025.

Table A.2: Robustness of Main IV estimates

Panel A. Full Sample

	Death Within 90 Days of Admission					
	(1)	(2)	(3)	(4)	(5)	(6)
Assigned No Roommate	-0.007**	-0.002	-0.008**	-0.009**	-0.016***	-0.014***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Assigned Roommate with ADRD	0.014***	0.012***	0.012***	0.012***	0.011***	0.011***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
F-statistic	25832	25798	18546	18554	22351	22336
Nursing Home × Year FE	X	X	X	X	X	X
Controls for Capacity Strain		X	X	X	X	X
County × Year × Month FE			X	X	X	X
Day of Week FE				X	X	X
Room FE					X	X
Controls for Baseline Characteristics						X
Dependent Variable Mean	0.155	0.155	0.155	0.155	0.155	0.155
Number of Observations	2,612,242	2,612,242	2,609,783	2,609,783	2,497,620	2,496,909

Panel B. Patients with AD/ADRD

	Death Within 90 Days of Admission					
	(1)	(2)	(3)	(4)	(5)	(6)
Assigned No Roommate	0.013**	0.019***	0.004	0.004	0.001	0.002
	(0.006)	(0.006)	(0.008)	(0.008)	(0.011)	(0.011)
Assigned Roommate with ADRD	0.051***	0.050***	0.047***	0.047***	0.051***	0.049***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.008)
F-statistic	10034	10035	5324	5325	5809	5802
Nursing Home × Year FE	X	X	X	X	X	X
Controls for Capacity Strain		X	X	X	X	X
County × Year × Month FE			X	X	X	X
Day of Week FE				X	X	X
Room FE					X	X
Controls for Baseline Characteristics						X
Dependent Variable Mean	0.161	0.161	0.161	0.161	0.161	0.161
Number of Observations	750,156	750,156	746,248	746,248	589,482	589,287

Panel C. Patients without AD/ADRD

	Death Within 90 Days of Admission					
	(1)	(2)	(3)	(4)	(5)	(6)
Assigned No Roommate	-0.014***	-0.010***	-0.015***	-0.015***	-0.023***	-0.021***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
Assigned Roommate with ADRD	-0.006	-0.007*	-0.006	-0.006	-0.010**	-0.015***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
F-statistic	17734	17724	13967	13972	13792	13795
Nursing Home × Year FE	X	X	X	X	X	X
Controls for Capacity Strain		X	X	X	X	X
County × Year × Month FE			X	X	X	X
Day of Week FE				X	X	X
Room FE					X	X
Controls for Baseline Characteristics						X
Dependent Variable Mean	0.144	0.144	0.144	0.144	0.144	0.144
Number of Observations	1,764,818	1,764,818	1,761,960	1,761,960	1,641,947	1,641,487

Notes: Controls for capacity strain are indicators for quartiles of occupancy rates residualized of nursing home x year x month fixed effects. Standard errors are clustered at the nursing home level.

Table A.3: OLS Estimates with Additional Controls

Panel A. Full Sample

	Death Within 90 Days of Admission					
	(1)	(2)	(3)	(4)	(5)	(6)
Assigned No Roommate	0.006***	0.006***	0.006***	0.006***	-0.005***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Assigned Roommate with ADRD	0.001	0.001	0.001	0.001	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Nursing Home × Year FE	X	X	X	X	X	X
Controls for Capacity Strain		X	X	X	X	X
County × Year × Month FE			X	X	X	X
Day of Week FE				X	X	X
Room FE					X	X
Controls for Baseline Characteristics						X
Dependent Variable Mean	0.155	0.155	0.155	0.155	0.155	0.155
Number of Observations	2,612,242	2,612,242	2,609,783	2,609,783	2,497,620	2,496,909

Panel B. Patients with AD/ADRD

	Death Within 90 Days of Admission					
	(1)	(2)	(3)	(4)	(5)	(6)
Assigned No Roommate	-0.000	0.000	0.001	0.001	0.004	0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Assigned Roommate with ADRD	-0.008***	-0.008***	-0.008***	-0.008***	0.022***	0.019***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Nursing Home × Year FE	X	X	X	X	X	X
Controls for Capacity Strain		X	X	X	X	X
County × Year × Month FE			X	X	X	X
Day of Week FE				X	X	X
Room FE					X	X
Controls for Baseline Characteristics						X
Dependent Variable Mean	0.161	0.161	0.161	0.161	0.161	0.161
Number of Observations	750,156	750,156	746,248	746,248	589,482	589,287

Panel C. Patients without AD/ADRD

	Death Within 90 Days of Admission					
	(1)	(2)	(3)	(4)	(5)	(6)
Assigned No Roommate	0.007***	0.007***	0.008***	0.008***	-0.007***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Assigned Roommate with ADRD	0.003***	0.003***	0.003***	0.003***	-0.002***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Nursing Home × Year FE	X	X	X	X	X	X
Controls for Capacity Strain		X	X	X	X	X
County × Year × Month FE			X	X	X	X
Day of Week FE				X	X	X
Room FE					X	X
Controls for Baseline Characteristics						X
Dependent Variable Mean	0.144	0.144	0.144	0.144	0.144	0.144
Number of Observations	1,764,818	1,764,818	1,761,960	1,761,960	1,641,947	1,641,487

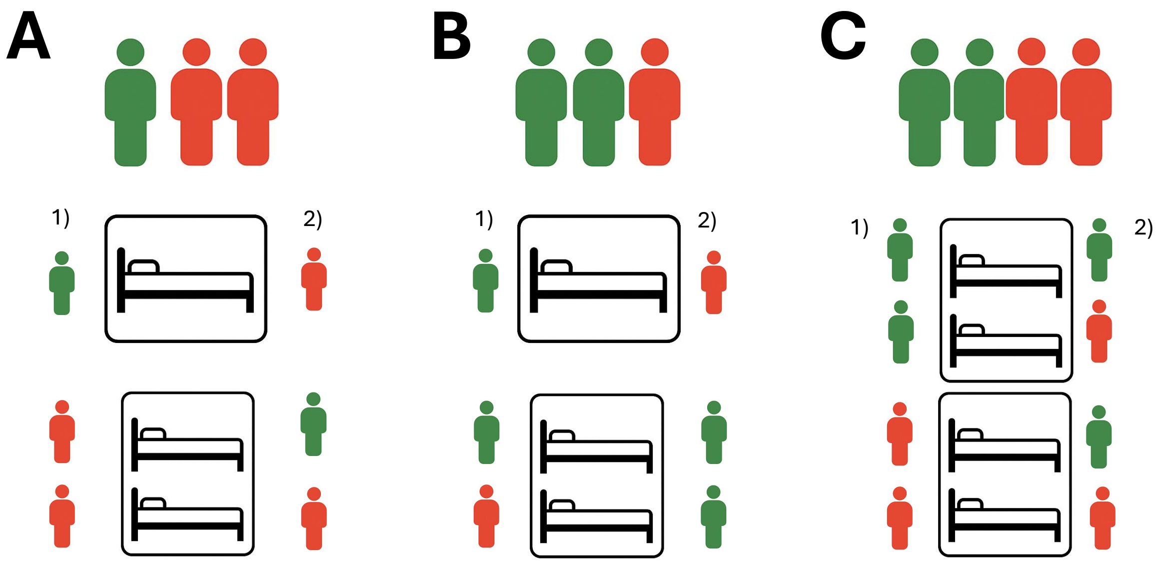
Notes: Controls for capacity strain are indicators for quartiles of occupancy rates residualized of nursing home x year x month fixed effects. Standard errors are clustered at the nursing home level.

Table A.4: Effect of Roommate Assignment on Mortality at Different Time Horizons

	Death Within 30 Days (1)	Death Within 90 Days (2)	Death Within 180 Days (3)	Death Within 360 Days (4)
Panel A. Full Sample				
Assigned to Room with No Roommate	-0.001	-0.007**	-0.009**	-0.006
	(0.002)	(0.003)	(0.004)	(0.004)
Assigned to Roommate with AD/ADRD	0.002	0.014***	0.014***	0.012***
	(0.002)	(0.003)	(0.004)	(0.004)
F-statistic	25,835	25,835	25,835	25,835
Dependent Variable Mean	0.063	0.155	0.226	0.312
Number of Observations	2,612,318	2,612,318	2,612,318	2,612,318
Panel B. Patients with AD/ADRD				
Assigned to Room with No Roommate	0.012***	0.013**	0.009	0.006
	(0.004)	(0.006)	(0.008)	(0.008)
Assigned to Roommate with AD/ADRD	0.018***	0.051***	0.062***	0.059***
	(0.004)	(0.006)	(0.007)	(0.008)
F-statistic	10,034	10,034	10,034	10,034
Dependent Variable Mean	0.058	0.161	0.247	0.357
Number of Observations	750,184	750,184	750,184	750,184
Panel C. Patients Without AD/ADRD				
Assigned to Room with No Roommate	-0.006***	-0.014***	-0.015***	-0.010**
	(0.002)	(0.004)	(0.004)	(0.005)
Assigned to Roommate with AD/ADRD	-0.007***	-0.006	-0.007	-0.006
	(0.003)	(0.004)	(0.005)	(0.005)
F-statistic	17,736	17,736	17,736	17,736
Dependent Variable Mean	0.056	0.144	0.210	0.286
Number of Observations	1,764,866	1,764,866	1,764,866	1,764,866

Notes: All regressions include nursing home-by-year fixed effects. Standard errors are clustered at the nursing home level.

Figure C.5: Room Assignments and 90-Day Mortality



90-Day Mortality	A1	A2	B1	B2	C1	C2
Average	16.8%	14.1%	13.7%	14.4%	16.7%	13.9%
AD/ADRD	18.5%	14.1%	13.4%	14.7%	18.5%	13.4%
No AD/ADRD	13.5%	14.3%	13.9%	14.3%	14.9%	14.3%

Notes: This figure illustrates the 90-day mortality rates by baseline health across different patient to room assignments. Panels A and B consider a single room (including multi-bed rooms with one assigned patient) and a two-bed room. Panel C considers two two-bed rooms. For each panel, we consider a census of patients shown at the top and illustrate two possible room assignments 1) and 2). Using the estimates from columns 1-2 in Panel A of Table 3, we calculate the corresponding average mortality rate and the mortality rates for patients with and without and AD/ADRD diagnosis at baseline and present them in the table below. (Column A1, for instance, corresponds to Panel A's allocation 1). Note that the 90-day mortality rate for the reference group , being assigned to a shared room, in which none of the roommates has a diagnosis of AD/ADRD, equals 13.4% for patients with AD/ADRD and 14.9% for patients without AD/ADRD. The mortality rate for patients without AD/ADRD in configuration A1 is then for example: 14.9%-1.4%=13.5%. The average mortality for patients with AD/ADRD is 13.4%+5.1%=18.5%. Average mortality is then (2*18.5%+13.5%)/3=16.8%.

AD/ADRD fares worse, large positive spillovers to patients with AD/ADRD produce a substantial net mortality reduction.

Example B considers the case of a healthy peer. In this case, average mortality is lower when assigning the single room to a patient without AD/ADRD. Finally, example C holds the room type fixed and isolates the interaction between patient and peer health. As discussed, patient and peer health are substitutes. Thus mixing patients by health reduces average mortality by 2.8 percentage points.

In conclusion, our estimates imply the different room assignments of patients have first order implications on average patient mortality. Average mortality varies by up to 2.8 percentage points. depending on the assignment. For comparison, Cheng (2023) estimates differences in value-added to 90-day mortality between nursing homes in California and finds a standard deviation of 2 percentage points. This suggests that differences in room assignments correspond to up to 2.8/2=1.4 standard deviations in quality between nursing homes. While 90-day mortality is almost certainly not the only outcome that nursing homes consider, our findings suggests that aggregate health outcomes can be very sensitive to room assignments.

D Optimal Room Assignment

D.1 Model for Optimal Assignment Rule

In this section, we present a simple model of an assignment rule that minimizes mortality, which we call an optimal assignment rule for simplicity (keeping in mind that in practice one may care about outcomes other than survival). We will then quantify the importance of room assignment for mortality, by comparing counterfactual mortality under this optimal assignment rule relative to alternative assignment rules.

Consider a static setting, where the social planner decides how to allocate patients with and without AD/ADRD to rooms in a given nursing home. Denote the fraction of patients with AD/ADRD by π and types by $k \in \{\text{AD}, \text{No}\}$. Letting N be the total number of patients and N_k be the number of patients of each type, we thus have:

$$N_{\text{AD}} = \pi N, \qquad N_{\text{No}} = (1-\pi)N.$$

Suppose the nursing home has R rooms, a fraction p of which are private, and that shared rooms each has a capacity of two. In addition, for simplicity assume that the total capacity of the nursing home is equal to the number of patients. Assuming that N is large so we can ignore integer constraints, this implies that:

$$N = pR + 2(1-p)R.$$

Note that if instead total capacity exceeds the number of patients (holding π and p constant), the extra degrees of freedom means that optimal assignment will generally be even more effective.

Denote the different types of room assignment — to a private room, a shared room with a roommate who has AD/ADRD, and a shared room with a roommate who does not have AD/ADRD — by $r \in \{0, \text{AD}, \text{No}\}$. We denote (average) potential outcomes for each type by:

$$Y_k(r) \equiv Pr\left(\text{Death Within 90 Days}|\text{Type} = k, \text{Assigned to Room of Type } r\right),$$

and let $u_{kr} \equiv 1 - Y_k(r)$ for convenience. The decision variables for the optimal assignment problem are:

- The number of AD patients in private rooms, x^0_{AD},
- The number of patients without AD in private rooms, x^0_{No},
- The number of shared rooms with two patients with AD, s_{AA},
- The number of shared rooms with two patients without AD, s_{NN}, and,
- The number of shared rooms with one patient who has AD and another who does not, s_{AN}.

The optimal assignment problem can thus be written as the following linear program:

$$
\begin{aligned}
\max_{x^0_{\text{AD}}, x^0_{\text{No}}, s_{\text{AA}}, s_{\text{NN}}, s_{\text{AN}}} \quad V = \quad & \underbrace{x^0_{\text{AD}} u_{\text{AD},0}}_{\text{AD in Private Rooms}} \quad + \quad \overbrace{x^0_{\text{No}} u_{\text{No},0}}^{\text{No AD in Private}} \\
& + \underbrace{2 s_{\text{AA}} u_{\text{AD,AD}}}_{\text{Shared: (AD,AD)}} + \overbrace{2 s_{\text{NN}} u_{\text{No,No}}}^{\text{Shared: (No, No)}} + \underbrace{s_{\text{AN}} \big(u_{\text{AD,No}} + u_{\text{No,AD}} \big)}_{\text{Shared: (No,AD)}}, \\
\text{s.t.} \quad & x^0_{\text{AD}} \geq 0, x^0_{\text{No}} \geq 0, s_{\text{AA}} \geq 0, s_{\text{NN}} \geq 0, s_{\text{AN}} \geq 0, \text{ (Non-Negativity)} \\
& x^0_{\text{AD}} + x^0_{\text{No}} = pR, \text{ (Balance for private rooms)} \\
& s_{\text{AA}} + s_{\text{NN}} + s_{\text{AN}} = (1-p)R, \text{ (Balance for shared rooms)} \\
& x^0_{\text{AD}} + 2 s_{\text{AA}} + s_{\text{AN}} = \pi N, \text{ (Balance for Patients with AD)} \\
& x^0_{\text{No}} + 2 s_{\text{NN}} + s_{\text{AN}} = (1-\pi) N, \text{ (Balance for Patients without AD)}
\end{aligned}
$$

where the second and third lines ensures that room assignments match the different types of rooms, and the last two lines ensures room assignments are consistent with the number of different types of patients. Note that one of the last four constraints is redundant, but we retain it for completeness.

Using the contraints and the fact that $x^0_{\text{No}} = pR - x^0_{\text{AD}}$, we can express s_{AA} and s_{NN} as:

$$
s_{\text{AA}} = \tfrac{1}{2}\big(\pi N - x^0_{\text{AD}} - s_{\text{AN}}\big), \qquad s_{\text{NN}} = \tfrac{1}{2}\big((1-\pi)N - (pR - x^0_{\text{AD}}) - s_{\text{AN}}\big).
$$

This simplifies the linear program to one with two choice variables, which (after some algebraic manipulation) we can write as:

$$
\begin{aligned}
\max_{x^0_{\text{AD}}, s_{\text{AN}}} \quad V = & C + \psi x^0_{\text{AD}} + \theta s_{\text{AN}} \\
\text{s.t.} \quad & x^0_{\text{AD}} \in [L, U], \\
& s_{\text{AN}} \in \big[0, \min\{\pi N - x^0_{\text{AD}}, (1-\pi)N - (pR - x^0_{\text{AD}})\}\big],
\end{aligned}
$$

where:

$$
\begin{aligned}
\psi &\equiv (u_{\text{AD},0} - u_{\text{AD,AD}}) - (u_{\text{No},0} - u_{\text{No,No}}), \\
\theta &\equiv (u_{\text{AD,No}} + u_{\text{No,AD}}) - (u_{\text{AD,AD}} + u_{\text{No,No}}), \\
C &\equiv \pi N u_{\text{AD,AD}} + (1-\pi) N u_{\text{No,No}} + pR(u_{\text{No},0} - u_{\text{No,No}}), \\
L &\equiv \max\{0, pR - (1-\pi)N\}, \quad U \equiv \min\{\pi N, pR\}.
\end{aligned}
$$

The parameter ψ can be interpreted as the difference in privacy premium between patients with and without AD. On the other hand, θ can be interpreted as the surplus from having a mixed room rather than homogeneous shared rooms. Moreover, we observe that θ is almost identical to the expression defining supermodularity or submodularity (between patients with and without AD/ADRD in shared rooms), with negative values of θ corresponding to supermodularity, and positive values corresponding to submodularity.

Given that the feasible set is a polytope, by the fundamental theorem of linear programming, an optimum lies at an extreme point. We consider different cases based on the sign of θ.

First, in the supermodular case ($\theta \leq 0$), homogeneous rooms is more efficient than mixed rooms, so we set $s^*_{\text{AN}} = 0$. In addition, we allocate as many private rooms to the type with the greater privacy premium as possible: if $\psi > 0$, set $x^{0*}_{\text{AD}} = U$, whereas if $\psi \leq 0$, set $x^{0*}_{\text{AD}} = L$.

On the other hand in the submodular case ($\theta > 0$), mixed rooms is more efficient than homogeneous rooms. Hence, we want to maximize the number of shared rooms, except when the difference in privacy premium outweighs the gains from mixing ($|\psi| > \theta$), in which case we may want to sacrifice some mixed rooms in order to assign the type that benefits disproportionately from the privacy premium to private rooms.

Start by considering the simple case where $\theta \geq |\psi|$. For a given choice of x^0_{AD}, the maximum feasible number of mixed rooms is:

$$s_{max}(x^0_{\text{AD}}) = \min\{ \overbrace{\pi N - x^0_{\text{AD}}}^{\text{Remaining Patients with AD}}, \underbrace{(1-\pi)N - (pR - x^0_{\text{AD}})}_{\text{Remaining Patients without AD}} \}.$$

This quantity is maximized when the two constraints bind simultaneously, which occurs when $x^0_{\text{AD}} = x_{bal}$:

$$x_{bal} \equiv \frac{(2\pi - 1)N + pR}{2}.$$

Hence, if $x_{bal} \in [L, U]$, we set $x^{0*}_{\text{AD}} = x_{bal}$ and $s^*_{\text{AN}} = s_{max}(x^{0*}_{\text{AD}})$. On the other hand, if $x_{bal} \notin [L, U]$, we set $x^{0*}_{\text{AD}} = \text{clip}(x_{\text{bal}}; [L, U])\}$ and $s^*_{\text{AN}} = s_{max}(x^{0*}_{\text{AD}})$, depending on which choice of x^{0*}_{AD} results in a larger V. If $|\psi| > \theta$, then we assign as many patients from the type that benefits more from the privacy premium to private rooms as possible: $x^{0*}_{\text{AD}} = \mathbb{I}[\psi < 0] \cdot L + \mathbb{I}[\psi \geq 0] \cdot U$ and $s^*_{\text{AN}} = s_{max}(x^{0*}_{\text{AD}})$.

D.2 Comparison of Counterfactual Mortality with Other Assignment Rules

To quantify potential gains from optimal assignment, we compare our optimal assignment rule to two other assignment rules — random assignment and perfect segregation.

Under random assignment, the fraction of patients of each type assigned to private rooms is equal to the fraction of total beds that are private:

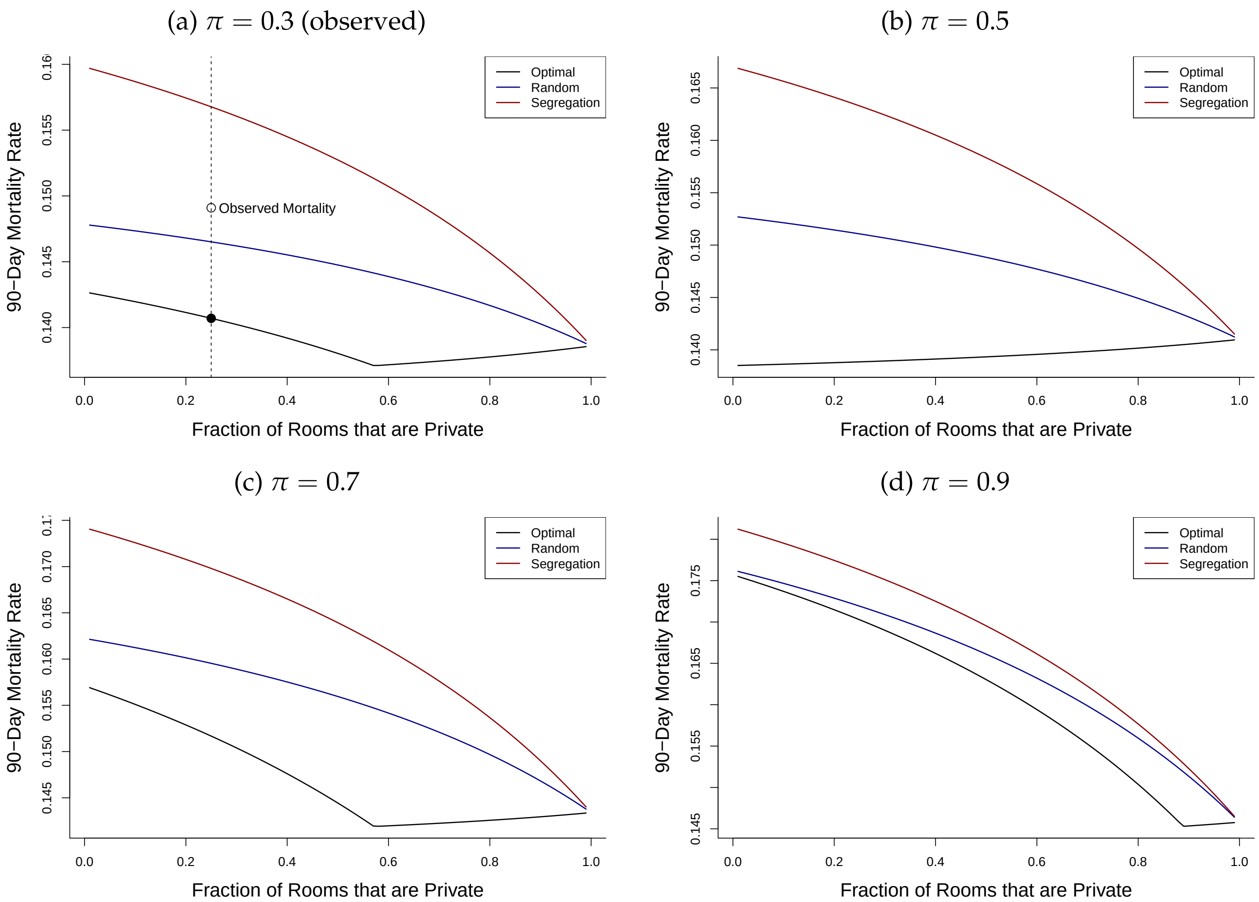
$$\frac{x^0_{\text{AD}}}{\pi N} = \frac{x^0_{\text{No}}}{(1-\pi)N} = \frac{p}{p + 2(1-p)}.$$

In addition, the fraction of shared rooms with two patients with (without) AD/ADRD is π^2 (respectively $(1-\pi)^2$), and the fraction of shared rooms that are mixed is $2\pi(1-\pi)$. Under counterfactual assignment with perfect segregation, we assume that private rooms are randomly assigned, but a fraction π of shared rooms have two patients with AD/ADRD and the remaining shared rooms have two patients without AD/ADRD.

Figure D.6 shows counterfactual 90-day mortality under different assignment rules, varying the proportion of the patient population having AD/ADRD (π) across different panels, and varying the proportion of rooms that are private in each panel. Average potential outcomes for different patient types under different room assignments $Y_k(r)$ are based on estimates in panel A of Table C.5. We observe that different assignment rules can lead to quantitatively meaningful differences in mortality: when half of the patient population has AD/ADRD and 10 percent of rooms are private, optimal assignment leads to a 0.7 percentage point fall in 90-day mortality relative to random assignment, and a 1.4 percentage point fall relative to perfect segregation.

In addition, we observe four qualitative patterns. First, for all values of π and p, mortality is highest under perfect segregation, followed by random assignment and finally (by defintion) optimal assignment. Second, mortality is decreasing in the share of rooms that are private, consistent with assignment to private rooms reducing mortality on average. Third, differences in counterfactual mortality tend to be greater when a roughly even share of the population has AD/ADRD, consistent with there being the most scope for assignment rules to make a difference when there is more heterogeneity in the population. Fourth, differences across different assignment rules are greatest when a small share of rooms are private, pointing to the importance of submodularity in health production across patient types in shared rooms.

Figure D.6: Counterfactual Mortality Under Different Assignment Rules



E Bounds Under Potential Violations of Conditional Independence

Let ρ denote the reduced form coefficients, and let Π denote the square 2×2 matrix of first stage coefficients. Then, the treatment effects β are given by:

$$\beta = \Pi^{-1}\rho.$$

Following Conley, Hansen, and Rossi (2012), we can write violations to conditional independence of the instruments as:

$$\hat{\rho} = \Pi\beta + \Delta,$$

so that the bias in the 2SLS estimate of β is given by:

$$\hat{\beta} - \beta = \Pi^{-1}\Delta.$$

The special case where $\Delta = 0$ corresponds to no violation of the conditional independence assumption, and thus no bias in the 2SLS estimate. In what follows, we will use methods from Oster (2019) to obtain bounds for Δ under different assumptions, and translate these bounds for the reduced form coefficients to bounds for the treatment effects.

Following Oster (2019), let R^2_{max} denote the R-squared from a hypothetical regression of the outcome on the instruments, as well as all observable and unobservables controls, and let δ denote the degree of selection on unobservables relative to selection on observables. In addition, denote the main reduced form estimates and the reduced form estimates with all controls by $\hat{\rho}_{main}$ and $\hat{\rho}_{rich}$ respectively, and similarly the R-squareds from these regressions by R^2_{main} and R^2_{rich} respectively. Then, the bias in the lth component of the reduced form estimate is bounded by:

$$|\hat{\rho}_{main} - \rho|_l \leq |\delta_l| \cdot |\hat{\rho}_{main} - \hat{\rho}_{rich}|_l \cdot \frac{R^2_{max} - R^2_{rich}}{R^2_{rich} - R^2_{main}}.$$

Denoting this bound by $|\Gamma|$, we can then bound the bias in the (lth component of the) 2SLS estimate by:

$$|\hat{\beta}_{main} - \beta|_l \leq (|\Pi^{-1}|_\circ |\Gamma|)_l,$$

where $|A|_\circ$ denotes the matrix A with absolute values applied entry-wise: $|A|_{\circ,ij} \equiv A_{ij}$.

Appendix Table E.5 shows these bounds for our IV estimates of treatment effects for the full sample, as well as for patients with AD/ADRD, and patients without AD/ADRD. The first two columns reproduces our treatment effect estimates with nursing home-by-year fixed effects or with a much richer set of fixed effects and controls for reference. Column 3 shows that under the values of δ and R^2_{max} recommended by Oster (2019)—equal selection ($\delta = 1$) and $R^2_{max} = 1.3R^2_{rich}$—the bounds do not include zero, suggesting that our IV estimates are robust to modest violations of conditional independence. Column 4 shows that under the extreme assumption that $R^2_{max} = 1$—i.e., that the instruments, observed controls, and unobserved controls can explain all the variation in 90-day mortality—most of the bounds include zero, although our results for the effect of being assigned a roommate with AD/ADRD for patients with AD/ADRD remain very robust. Our IV estimates may also be even more robust to violations of conditional independence than these bounds suggest for another reason: for several estimates the addition of controls strengthens our results, suggesting that violations of the conditional independence assumption may bias us against finding an effect.

Alternatively, rather than assume values for R^2_{max} and δ and then obtain bounds for β, for a given value of R^2_{max} we can back out the minimum δ in order for violations of conditional independence to explain our non-zero treatment effect estimate (i.e., such that the bounds contain zero). Hence, we can trace out an "indifference curve" in (R^2_{max}, δ) space, such that our treatment effect estimates are robust (respectively, not robust) to violations of conditional independence for values of (R^2_{max}, δ) to the southwest (northeast) of this curve.

Appendix Figure E.7 plots these curves, with each panel corresponding to a treatment effect for a particular sample. For example, panel (a) corresponds to sensitivity analysis for the average treatment effect of being assigned to a no roommate (relative to being assigned roommates without AD/ADRD) in the full sample, while panel (c) corresponds to sensitivity analysis for the treatment effect for patients with AD/ADRD of being assigned a roommate also with AD/ADRD (relative to being assigned roommates without AD/ADRD). Similar to bounds in Appendix Table E.5, we observe that our results on the effect of being assigned a roommate with AD/ADRD for individuals with AD/ADRD are the most robust: if we assume $R^2_{max} = 0.5$ (respectively, $R^2_{max} = 1$), selection on unobservables have to be more than 100 (10) times more important than selection on observables in order to explain the positive IV estimate.

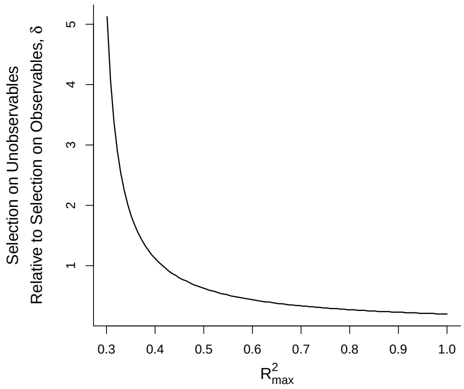
Table E.5: IV bounds Under Potential Violations of Conditional Independence

	Baseline Estimate	Controlled Estimate	Bounds with $\delta = 1$, $R^2_{\max} = 1.3R^2_{rich}$	Bounds with $\delta = 1$, $R^2_{\max} = 1$
Panel A. Full Sample				
	(1)	(2)	(3)	(4)
Assigned to Room with No Roommate	-0.007** (0.003)	-0.014*** (0.003)	[-0.0104, -0.0027]	[-0.0407, 0.0275]
Assigned to Roommate with AD/ADRD	0.014*** (0.003)	0.011*** (0.004)	[0.0121, 0.0153]	[-0.0005, 0.0279]
Panel B. Patients with AD/ADRD				
	(1)	(2)	(3)	(4)
Assigned to Room with No Roommate	0.013** (0.006)	0.002 (0.011)	[0.0061, 0.0198]	[-0.0136, 0.0395]
Assigned to Roommate with AD/ADRD	0.051*** (0.006)	0.049*** (0.008)	[0.0502, 0.0528]	[0.0464, 0.0565]
Panel C. Patients without AD/ADRD				
	(1)	(2)	(3)	(4)
Assigned to Room with No Roommate	-0.014*** (0.004)	-0.021*** (0.004)	[-0.0184, -0.0087]	[-0.0479, 0.0208]
Assigned to Roommate with AD/ADRD	-0.006 (0.004)	-0.015*** (0.005)	[-0.0101, -0.0024]	[-0.0337, 0.0211]

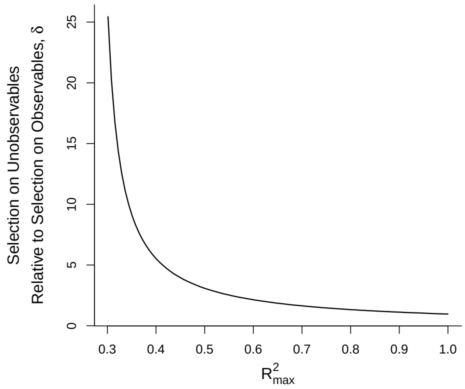
Notes: This table shows bounds for the treatment effect estimates allowing for different amounts of selection on unobservables relative to observables. Columns (1) and (2) report the baseline and controlled estimates, respectively. Columns (3) and (4) show bounds under alternative assumptions about the amount of variation in the outcome that can be explained by instruments, observed controls, and unobserved controls, as well as the maximum R^2 attributable to unobserved controls to explain all of the variation in the outcome.

Figure E.7: Values of $R^2_{\max}$ and δ Required to Explain Treatment Effect Estimates

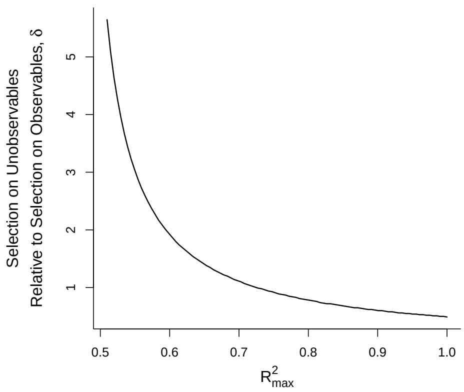
Full Sample
(a) (Empty Room)



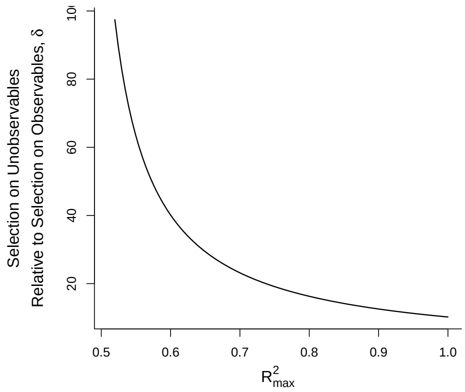
Full Sample
(b) (Roommate with ADRD)



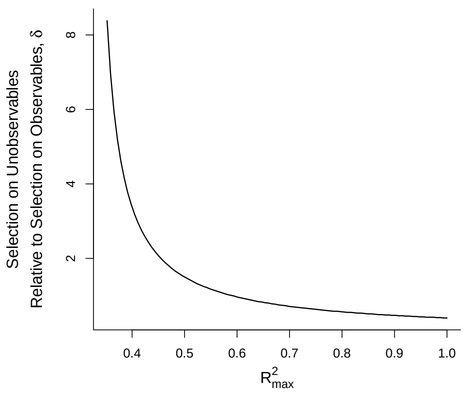
ADRD Sample
(c) (Empty Room)



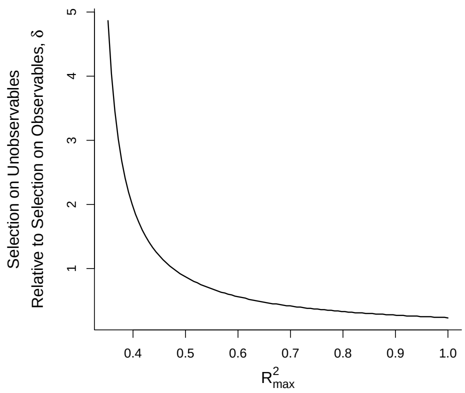
ADRD Sample
(d) (Roommate with ADRD)



Non-ADRD Sample
(e) (Empty Room)



Non-ADRD Sample
(f) (Roommate with ADRD)



Notes: Each panel of this figure plots values of (R^2_{max}, δ) under which one of the endpoints of the bounds for the IV estimates contains zero. Each row—e.g., panels (a) and (b)—shows the robustness of the two IV estimates (for the effect of being assigned no roommate or the effect of being assigned a roommate with AD/ADRD) for a particular sample (either the full sample, the subsample of patients with AD/ADRD, or the subsample of patients without AD/ADRD).