# Vertical Integration and Agency with Cognitive Limitations: Evidence from Long-Term Care*

Toshiaki Iizuka† Shinya Sugawara‡

The University of Tokyo Tokyo University of Science

Masaki Takahashi§

Sophia University

*Work in progress. Not for quote.*

This version: November 16, 2024

**Abstract**

We investigate vertical integration and agency in at-home long-term care (LTC) services in Japan, where a care manager (CM) works as an agent to coordinate care for frail older people who often have cognitive constraints. Importantly, CMs and downstream services can legally be held in the same ownership in Japan. This raises the concern that financial incentives from vertical integration may distort a CM's decision as an agent for older people. We find that CMs play key roles in determining LTC services for frail older people. Vertically integrated CMs induce demand for co-owned providers by creating a care plan that includes service types provided by co-owned services, especially for recipients with higher cognitive limitations. Moreover, integrated CMs substitute not-co-owned services with co-owned services. However, we find no evidence that vertical integration improves efficiency by slowing the progression of care needs or deterioration of health outcomes. Thus, our results indicate that vertical integration of CMs and service providers leads to an agency problem by distorting care plans in favor of integrated services without clear benefits to recipients, and the concern for agency conflicts is greater when recipients have more cognitive limitations.

**Keywords**: long-term care, vertical integration, principal-agent model, physician agency, supplier-induced demand, boundaries of the firm, care coordination

**JEL codes**: I11, I18, J14, L14

---

* We gratefully acknowledge financial support from JSPS KAKENHI Grant Number 18H0086 and 21H00722. The authors thank Chie Hanaoka, Momotazur Rahman, and seminar participants at AEA and JHEA Meetings 2024. All remaining errors are our own.

† Graduate School of Economics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, JAPAN. Email: toshi.iizuka@gmail.com

‡School of Management, Tokyo University of Science, 1-11-2, Fujimi, Chiyoda-ku, Tokyo, 102-0071, JAPAN. Email: shinya_sugawara@rs.tus.ac.jp

§ Department of Economics, Sophia University, 7-1 Kioicho, Chiyoda-ku, Tokyo, 102-8554, JAPAN. Email: masaki-takahashi@sophia.ac.jp

# 1. Introduction

In many healthcare services, recipients often lack the knowledge or capability to choose the best service and rely on a knowledgeable agent for a recommendation. For example, patients rely on a primary care physician to recommend a hospital for heart surgery and a cardiac surgeon for specific surgical procedures. One concern in these situations is that when the agent and downstream services are under the same ownership, the agent's recommendation may be influenced by the financial incentive to benefit co-owned services. That is, integrated agents may not act in the client's best interests. On the other hand, vertical integration may help the client. For example, vertical integration may improve efficiency by accelerating information flow within the boundaries of the firm, which may reduce duplicated activities, such as tests and treatments, and improve health outcomes. Therefore, whether vertical integration is beneficial to patients is an empirical question.

We investigate these issues in the context of long-term care (LTC) services in Japan, where a care manager (CM hereafter) works as an agent for frail older people and coordinates care on their behalf. Choosing the best LTC service is difficult for frail older people who often have cognitive limitations. Instead, a CM creates a care plan for them, determining the type of service, service volume, and service provider. Importantly, CMs and downstream services can legally be held in the same ownership in Japan. This raises the concern that financial incentives from vertical integration may distort the CM's decision as an agent for older people, especially when recipients have cognitive limitations. In fact, numerous anecdotes exist that vertically integrated CMs are under considerable pressure to create a care plan that increases the profits of LTC providers under the same ownership (MHLW 2004). Although who coordinates care may vary from country to country, frail older people must rely on an agent to choose the best LTC service for them in any setup. Thus, the potential agency problem in LTC services due to vertical integration is a universal issue in all aging societies.

In this paper, we study the demand-inducement and efficiency effects of vertical integration in at-home LTC services. For demand inducement effects, we analyze whether vertically integrated CMs induce demand to benefit co-owned downstream services regarding i) the type of service included in a care plan ("service choice") and ii) the volume of service that the client receives ("service volume"). Additionally, we examine "service substitution," i.e., whether integrated CMs substitute not-co-owned services with co-owned services. For the analysis of efficiency effects, we study two sets of outcomes. First, we examine whether vertical integration slows the progression of care needs regarding the minutes required to care for the recipient. Second, we study the impact of vertical integration on health outcomes, using limitations in activities of daily living (ADLs), instrumental activities of daily living (IADLs), and dementia dependency, the level of dependency in daily living for people with dementia. Note that care needs and health outcomes are closely related but different concepts. Care needs represent the level of care required, while ADL/IADL limitations and dementia dependency

capture the level of frailty of older people. Notably, the level of care needs is not a monotone transformation of ADL/IADL limitations. For example, as the Ministry of Health, Labour and Welfare (MHLW) emphasizes, caring for a healthy recipient with dementia may require more time and effort than caring for a bedridden dementia recipient.[1] Therefore, it is important to look at the impact of vertical integration on care needs and health outcomes separately.

We use three sets of data in the analysis. The first is individual-level monthly administrative claims data that record item-by-item LTC service use and expenditure covered by the pubic LTC insurance (LTCI). This data also includes information on the timing of (re)certification, the required level of care (care level), and the characteristics of the user, such as age, gender, and coinsurance rate. The second data set is certification data, including each recipient's certification information, such as care minutes. The data also includes measures for ADL, IADL, and dementia dependency. The certification information is updated whenever the individual is recertified. We obtained the above data from one of the top 20 municipalities in Japan by population, with a population of around 1 million. The two data sets can be linked using anonymized identification numbers. The third data set contains ownership information of CMs and service providers. We use this data set to determine which CMs and service providers have the same ownership, i.e., vertically integrated. Our data include recipients whose baseline certification occurred between June 2012 and November 2018.

Throughout the paper, we focus on at-home LTC services that provide care for those who stay at their own homes instead of at institutions, such as nursing homes. Our CMs coordinate care only for recipients who receive at-home LTC services. Formal at-home LTC services have grown enormously as insurers in many countries seek to care for recipients at home instead of institutions (De Meijer et al., 2015). We focus on three primary at-home LTC services, i.e., adult daycare, outpatient rehabilitation (rehabilitation hereafter), and home care, which account for more than 71% of all non-institutional formal LTC expenditures in our data.

Our primary identification concern is the selection of a CM by recipients. The threat is that unobserved user preferences may be correlated with both the vertical integration status of the CM and the outcomes of our interest. For example, recipients who prefer a particular service, say rehabilitation service, may choose a CM who vertically integrates rehabilitation service. Alternatively, recipients who strongly prefer a specific daycare provider may not select an integrated CM, knowing that the integrated CM will refer to a co-owned daycare provider. We address these endogeneity issues using a "differential distance (DD)" instrument. Specifically, we calculate the distance from the user's residence to the closest CM, vertically integrated and not integrated, and take the difference. We then use it as an instrument for vertical integration. We calculate the DD instrument for each service as CMs do not always integrate all types of downstream services. We expect the shorter the relative distance to an integrated CM, the more

[1] MHLW discusses this possibility on its website https://www.mhlw.go.jp/topics/kaigo/nintei/gaiyo2.html (In Japanese. Accessed 2023.6.10)

likely the user will choose an integrated CM. The instrument is valid if this relationship holds and individuals do not choose a residence based on the distance to CMs. We confirm that the first "relevance" relationship holds in our data. Although we cannot directly test the latter assumption, we show that user attributes are balanced by the differential distance measure, supporting the assumption. The DD instrument has been used in previous studies of healthcare services (e.g., McClellan et al. 1994; Grabowski et al. 2013; Rahman et al. 2016; Konetzka et al. 2018; Card et al. 2023) and works well in our case, too.

We also consider an additional endogeneity issue for the analysis of efficiency effects. In particular, we are concerned that other unobserved CM attributes correlated with the vertical integration status may independently affect the outcome of our interest. For example, if vertically integrated CMs provide higher-quality services, it may be difficult to distinguish vertical integration's impact from unobserved quality on our health outcome measures.[2] Similar to Card et al. (2023), we address this issue by including a dummy variable that captures the quality of the CM as an additional endogenous regressor. Specifically, we assign a high- versus low-quality status for each CM based on their recipients' average improvements in care minutes since the baseline certification. To address the endogeneity of the quality status, we use an additional DD instrument similar to the previous ones, which calculates the relative distance to the closest high- versus low-quality CMs.

We find evidence that vertically integrated CMs induce demand for co-owned downstream providers. First, regarding service choice, we find that vertically integrated CMs are more likely to include co-owned service types in a care plan. For example, if a CM and a daycare service are co-owned, the probability of including daycare in a care plan increases by more than 40%. Importantly, this impact of vertical integration on service choice is most visible for recipients with higher cognitive limitations. We also find suggestive evidence for service substitution; integrated CMs are less likely to include not-co-owned services in a care plan, although the estimates are mostly insignificant.

Second, we find similar results to service volume. Vertical integration substantially increases the volume of downstream services CMs co-own while reducing not-vertically integrated services. Interestingly, examining the impact of vertical integration on the sum of the three service expenditures, we find no evidence that total expenditure increases with vertical integration. This result further indicates that vertical integration's primary effect is substituting not-co-owned services with co-owned services. CMs may substitute services instead of increasing total expenditure because there is a maximum amount of spending that public LTC insurance covers. Alternatively, recipients may have time constraints; receiving co-owned services may force the user to cut back on receiving other services. We continue to find that the volume inducement effect is larger for recipients with more cognitive limitations.

[2] We failed to find similar endogenous variables in the analysis of demand-inducement effects.

The results so far suggest that vertically integrated CMs' service choices may be suboptimal as they recommend co-owned service types at the expense of not-co-owned services. It is also concerning that the demand-inducement effect is greater for recipients with more cognitive limitations. However, we cannot yet conclude that these results indicate agency problems because these seemingly suboptimal choices may be compensated by efficiency gains owing to vertical integration, such as better coordination and information flow within the boundaries of the firm, which we investigate next.

Third, in contrast to the inducement effects we find above, we find no evidence supporting the efficiency effect of vertical integration. Specifically, we did not find that vertical integration slows the progression of care needs regarding required care minutes per day one or two years after the baseline certification. Results for health outcomes are similar; vertical integration does not help delay the deterioration of health outcomes regarding ADL/IADL limitations and dementia dependency.

In summary, vertical integration induces demand in the form of service substitution that would benefit co-owned downstream services. However, we do not find evidence for efficiency gain owing to the substitution, at least based on the measures we examined. These results indicate that vertical integration of CMs and LTC services leads to an agency problem by distorting care plans in favor of integrated services without clear benefits to recipients.

This study contributes to two strands of literature. The first is literature that examines how vertical integration affects agents' referral decisions. In healthcare services, studies have shown that vertically integrated agents refer patients to co-owned hospitals (Nakamura et al. 2007; Baker et al. 2016; Ho and Pakes 2014), refer more profitable users to own skilled nursing facilities (SNFs) (Cutler et al. 2020), perform more cardiac surgery (Afendulis and Kessler 2007), and prescribe more profitable drugs to patients (Iizuka 2007, 2012), among others. We contribute to the literature in three ways. First and most importantly, we study agents' referral decisions when their clients have cognitive constraints. Although the principal-agent relationship under cognitive constraints is an important area that economics can contribute (Chandra et al. 2023), to our knowledge, little research exists on this issue. Comparing recipients with different levels of dementia, we find that the demand-inducement effect is more significant for a higher level of dementia. Second, we study CMs, agents who play crucial roles in coordinating non-institutional LTC services, a sector becoming increasingly important in all aging societies. CMs play key roles because many alternative services are available, and frail older people must rely on an agent to choose the appropriate one. To our knowledge, Kishida (2015) is the only study that examines the impact of vertical integration of CMs on service choice (but not service volume and efficiency). However, it uses small survey data and does not fully address the endogeneity of vertical integration.[3] Third, we add to the literature by

[3] Sugawara and Nakamura (2016) is also related to our study. It examines the relationship between the density of vertically integrated CMs and LTC expenditures using municipality-level data from Japan,

studying agent "service choice" instead of "destination choice." Most previous studies examine agents' decision to refer to a co-owned or not-co-owned provider, i.e., the service's destination, conditional on a service such as hospital care or SNFs.[4] However, before deciding on a destination, agents often must choose a service from alternative service types, such as hospital care, SNFs, and home care. We contribute to the literature by studying service choice, an additional layer that may also be influenced by financial incentives created by vertical integration. Studying service choice is important because vertical integration may result in the service not being optimal for the recipient. This, in fact, could be more problematic than distorted destinations because, in the latter case, the same service will be provided regardless of the destination.

Our study also contributes to the literature on the effect of vertical integration in healthcare services. This literature is rapidly growing because vertical integration of healthcare services has substantially increased in the United States in recent years (Cutler et al., 2020). Unlike the first literature, this literature does not explicitly study agents' referral decisions. Instead, they study the consequences of vertical integration in health care on various outcomes, including healthcare utilization and expenditures (Baker et al. 2014; Koch et al. 2017; Lin et al. 2022), health outcomes (Koch et al. 2021), and price (Baker et al. 2014), among others. Most of these studies examine physician-hospital integration in the U.S., but several others have looked at integration between hospitals and skilled nursing facilities (SNFs) in the U.S. (e.g., Rahman et al. 2016, Konetza et al. 2018). We contribute to the literature by studying the consequences of vertical integration of care coordinators and providers in non-institutional LTC services, which has not been analyzed. Additionally, we add to the literature by proposing new efficiency measures of vertical integration, i.e., the level of care needs (required care minutes) and health outcomes (ADL/IADL limitations and dementia dependency) relevant to LTC services. Most previous studies use mortality and hospital readmission to measure health outcomes. Although highly relevant, these events are relatively rare, and thus, precisely estimating the efficiency effects is often difficult. In contrast, our measures vary more by frail older people, which could help detect the efficiency effects of vertical integration.

## 2. Background

### 2.1. Long-term care insurance in Japan

The Japanese government introduced the LTCI in 2000 due to rapid population aging. The LTCI is mandatory social insurance with universal coverage for all certified adults 65 years old or older or those with specific aging-related diseases from 40 to 64 years old. In 2021,

finding a positive correlation. Although they do not directly examine agent behavior, the results are consistent with demand inducement by CMs.

[4] Exceptions include Afendulis and Kessler (2007) and Iizuka (2007), which examine the impact of vertical integration on the choice of alternative cardiac treatments and hypertension drugs, respectively.

approximately 6.9 million people used the LTCI service in a month, and the total expenditure was 11.3 trillion JPY (or 102.4 billion USD[5]) or 2.08% of GDP.[6] The LTCI provides only in-kind benefits and has no cash payment for informal caregiving. A single municipality or a federation of several municipalities insures LTCI.

The LTCI covers broad services, including both at-home and institutional care. This study focuses on formal at-home LTC services received by older people living in their own homes. Providing care at home instead of at a nursing home has been a major policy drive in Japan and elsewhere,[7] as institutional care is believed to be more expensive.[8] Accordingly, the share of at-home LTC expenditure in Japan has steadily increased over time from 37% in 2001 to 49% in 2021.[9] Both for-profit and not-for-profit firms provide at-home services.

We focus our attention on three major at-home services: (1) adult daycare, (2) home care, and (3) rehabilitation, which comprise more than 71% of at-home care expenditures in our data. Another major at-home service is short stay, which accounts for about 15% of total at-home care spending. We did not include it as part of our analysis for two reasons. First, a short stay is primarily intended to provide "respite" for family caregivers – which allows caregivers to get away from caregiving and refresh themselves physically and mentally – and is, therefore, not likely to affect care recipients' care needs and health outcomes. Second, recipients use the three services regularly, usually a few times weekly. In contrast, a short stay is used non-regularly as an add-on service, especially when family members need the service. Thus, CMs are less likely to substitute a short stay with other regular services.

LTC service providers are paid for fees-for-services except for CMs. CMs are paid by capitation, with a fixed monthly fee (10,000 JPY or 91 USD) per recipient, regardless of the content or service volume of the care plan. Although CMs are paid by capitation, numerous anecdotes exist that vertically integrated CMs are under considerable pressure to create a care plan that increases the profits of LTC providers under the same ownership (MHLW 2004). Most recipients pay the coinsurance rate of 10% up to the maximum amount covered by the insurance, which differs by the care-needs level. Higher-income individuals pay 20% or 30%, while those

---

[5] Throughout this research, we use the average exchange rate in 2018, 1USD = 110.40 JPY, taken from the Bank of Japan homepage (https://www.boj.or.jp/statistics/market/forex/fxdaily/ex2018.pdf ).

[6] Source: Annual Report on Long-Term Care Insurance by the Ministry of Health, Labor and Welfare for LTCI statistics (https://www.mhlw.go.jp/topics/kaigo/osirase/jigyo/21/index.html), GDP Statistics by the Cabinet Office (https://www.esri.cao.go.jp/jp/sna/menu.html )

[7] Many countries have recently been exploring alternative solutions to reduce the expenses associated with institutional care. Formal care at home has emerged as a practical substitute, as suggested by Carmichael and Sevak (2005). In the Netherlands, the transition to formal care at home is regarded as the primary factor contributing to the decrease in the utilization of institutional care, as observed by De Meijer et al. (2015).

[8] Bakx et al. (2020) shows that, however, it may not be the case when one takes healthcare service expenditures into account.

[9] Source: The Survey of Long-Term Care Benefit Expenditures conducted by the Ministry of Health, Labour and Welfare. The numbers are taken from Kosei rodo tokei kyokai (2011) for 2001 and from https://www.mhlw.go.jp/toukei/saikin/hw/kaigo/kyufu/21/dl/03.pdf for 2021.

under social welfare pay nothing.

We briefly describe the three services we examine. Adult daycare is a service where care recipients stay in a facility during the daytime. Adult daycare aims to reduce the isolation of users confined to their homes, maintain their physical and mental functions, and reduce the burden of family caregivers. The facilities provide daily living support such as meals, bathing, and functional training to improve daily living functions. There is also interaction among older people, such as group activities to enhance daily living functions. Outpatient rehabilitation is also provided at a facility during the day. The service offers daily living support, such as meals and bathing, and functional training to improve daily living and oral function. The latter function is more primary in outpatient rehabilitation. Daycare and outpatient rehabilitation generally include pickup services. Home care is a service where professional caregivers visit the homes of users. The service provides nursing care such as meals, toileting, and bathing (physical care) and assistance in daily living such as cleaning, laundry, shopping, and cooking (daily living assistance) so that users can lead independent daily lives at home.

### 2.2. Certification process

One must be certified and assigned to a care-needs level to use public LTC services. A care-needs level reflects the level of care required, which determines LTC services recipients can use, the unit price for those services, and the maximum spending covered by LTCI. There are seven care-needs levels, consisting of two broad categories, "assistance required" (Yoshien, hereafter AR) and "care required" (Yokaigo, hereafter CR). AR and CR categories consist of two and five levels, respectively. In this paper, we focus on CR-level recipients.

The care-needs level is certified as follows.[10] When an eligible person applies for LTCI, a municipal investigator visits the person and fills out a questionnaire consisting of 74 items on health status. Based on the answers to these questions, a computer algorithm calculates "care-required minutes (care minutes hereafter)," estimated minutes required to provide LTC per day, and based on the care minutes, a care level is tentatively assigned to the individual. For example, those with care minutes from 50 to 70 and 70 to 90 are set to CR2 and CR3, respectively. Then, a municipal certification committee of several healthcare professionals meets and decides the care level by considering the tentative care level and assessments by the recipient's primary doctor and the municipal investigator.

In addition to the 74 questions determining the care minutes, municipal investigators assess the level of dependency in daily living for people with dementia ("dementia dependency" hereafter) during the certification interview. Dementia dependency has six levels, starting from Level 0 (no dementia) to 5 (most severe).[11] For example, Level 3 recipients sometimes have

[10] See Tsutsui and Muramatsu (2005) and Fu et al. (2023) for additional institutional details.
[11] Officially, the MHLW calls the latter “Level M” instead of Level 5.

symptoms and behaviors that interfere with daily life, have difficulty communicating, and require nursing care. Hisano (2009) shows that dementia dependency is highly correlated with the level of cognitive impairment assessed by the Mini-Mental State Examination (MMSE). Thus, dementia dependency can be considered a proxy measure of cognitive impairment due to dementia.

Care recipients must receive recertification before the expiration date determined by the certification committee. They can also request recertification by themselves when, for example, their health status worsens. Most recipients in our data receive recertification every year. Importantly, all the information required to determine the care level, including the 74 items on health status and dementia dependency, are re-assessed during recertification, which we use to assess the impact of vertical integration on care needs and health outcomes.

### 2.3. CM and Vertical Integration

CMs (or care coordinators) play a vital role in the Japanese LTCI. There are many at-home care services, and choosing the best one to serve their care needs is challenging for frail older people who often have cognitive limitations. CMs act as agents for the recipient and create a care plan on their behalf, choosing services, providers, and the volume of the service. There is no specific rule on choosing a CM, but most insurers (municipalities) provide a list of care management offices, and care recipients often select one from the list. CMs visit the recipient's home to assess their health status and care needs and create a care plan. In some cases, they visit service providers together and decide on one. After creating a care plan, the CM must visit the recipient at least once a month to monitor and update the care plan if necessary. Recipients can also create a care plan alone, but it is rare.[12] Thus, we focus on recipients whose CMs made care plans on their behalf in this study.

As noted in the Introduction, there is a long-standing concern that vertical integration of healthcare providers may create an agency problem. In the context of Japanese LTC services, vertically integrated CMs may make care plans and referral decisions that benefit co-owned downstream firms at the expense of their clients. Numerous anecdotes indicate that vertically integrated CMs are under considerable pressure to create care plans that include services provided by vertically integrated downstream providers, as noted by more than 200 CMs interviewed by the MHLW (MHLW 2004). Although the Japanese government was aware of such a possibility,[13] it did not bun vertical integration. Instead, the Ministry created a rule that may reduce self-referrals; the government reduced the capitation rate for CMs by 20% if a care

[12] According to Mitsubishi Sougou Kenkyuujo (2020), only 0.05% of care plans were created by recipients themselves in 2019.
[13] Please see, for example, the documents for the Social Security Council on October 4, 2005, published by the Japanese Ministry of Health, Labor and Welfare, https://www.mhlw.go.jp/shingi/2005/10/s1004-4.html.

management office referred more than 80% of recipients to a specific service provider.[14] [15] However, concerns remain that vertically integrated CMs may not be impartial and neutral. There have been repeated discussions on ensuring the fairness and neutrality of care management.[16]

## 3. Data Description

**Data Source** Our analysis utilizes three datasets. The first is LTCI administrative claims data, which contains individual-level monthly information on item-by-item LTC service use and expenditures. The data also includes the recipient's care needs level and characteristics, such as age, gender, coinsurance rate, and the identification number for the CM and service providers. The second data set is LTCI certification data that contains the recipient's care needs level, care minutes, responses to 74-item questions related to ADL/IADL limitations, and the municipal investigator's assessment of dementia dependency. These certification data are available when the recipient receives initial and subsequent certification. The certification information is updated whenever the individual is recertified. Another vital information in this dataset is the 7-digit postal code[17] of each recipient's address, which we use to calculate the distance from the recipient's residence to that of the CM and service providers.[18] We obtained the above two datasets from one of the top 20 municipalities in Japan by population, which has a population of around 1 million. The data covers recipients whose baseline certification occurred between June 2012 and November 2018.

The third dataset contains ownership information of CMs and downstream service providers, i.e., who owns them.[19] We use this information to determine which CMs and service providers are vertically integrated. Additionally, the data provides information on providers' physical location and characteristics, such as the number of CMs in the care management office. Again, we use the location information to measure the distance between the user, the CM, and the service provider. We obtained the data for three years, i.e., 2014, 2017, and 2020, and

[14] To be precise, this rule was applied to adult daycare, home health care, and equipment rental.
[15] In May 2015, the government penalized 7.6% of care management offices in Japan for this reason. Source: Report on Social Security Council (7/5/2015), the Ministry of Health, Labor and Welfare, https://www.mhlw.go.jp/file/05-Shingikai-12601000-Seisakutoukatsukan-Sanjikanshitsu_Shakaihoshoutantou/0000170287.pdf p.21
[16] Source: https://www.ihep.jp/wp-content/uploads/2020/04/19314.pdf in Japanese. (Accessed 2024/07/22).
[17] The seven-digit postal codes cover on average 1.13km$^2$ or 280 acres in our data.
[18] For the address of individual users, we use the central point of the postal code. If a user and a provider are located in the same postal code, we define the distance between them is zero. Otherwise, a distance between two points is measured as R arccos[sin $\theta_1$ sin $\theta_2$ + cos $\theta_1$ cos $\theta_2$ cos( $\gamma_1 - \gamma_2$)], where $\theta_1$ and $\theta_2$ are the latitudes in radians of the two points, while $\gamma_1$ and $\gamma_2$ are the radian longitudes. R denotes the radius of the earth for which we adopt R = 6,371 km as in Banerjee et al. (2004, p. 19).
[19] We purchased the data set, PAREA Care-II, from Kokusai Kogyo, Co., LTD (https://biz.kkc.co.jp/en/ (accessed 2023.07.09)

merged them with the claims data for 2012-2014, 2015-2017, and 2018-2020, respectively.

**Analysis Sample** We construct the analysis sample in the following way. For each recipient, we call the first certification that appeared in the data after June 2012 "baseline certification." We have approximately 35,000 baseline certifications in the data. For the "service choice" analysis, we focus on recipients who have been certified for at least 12 months since the baseline certification. This means, for example, those who were not recertified after six months are not included in the data. Later, in Section 6.1, we check whether the attrition rate differs by the vertical integration status of the CM, finding no such evidence. For the "service volume" analysis, similarly, we include individuals who were certified at least 12 months (for the short-term analysis) and 24 months (for the medium-term analysis) since the baseline certification.

For efficiency effects, we include individuals who are certified as of the $14^{th}$ month (for the short-term analysis) and the $26^{th}$ month (for the medium-term analysis) after baseline certification. We chose to investigate the $14^{th}$ month because the length of certification is typically one year, and thus, certifications expire no later than the $13^{th}$ month after the baseline certification. Then, the outcome variables of our interest, i.e., care minutes, ADL and IADL limitation, and dementia dependency, are updated at least once before the $14^{th}$ month; by using the outcome variables as of the $14^{th}$ month, we can analyze how vertical integration affects these outcomes. Again, we check whether the attrition rates – those not recertified as of the $14^{th}$ and $26^{th}$ month – differ by the vertical integration status in Section 6.1, finding no such evidence.

**ADL/IADL limitation score**

To measure older people's frailty, we construct the ADL/IADL limitation scores using questions asked by municipal investigators during the certification process. For the ADL limitation score, we focus on seven ADL-related questions (ability to use the toilet, transfer from chair, transfer from bed, eat meals, dress, bathe, and go out) and create a score similar to the Barthel Index (Mahoney and Barthel, 1965), a leading index of ADL limitations. Specifically, we create a dummy variable for each activity that equals one if the recipient has a limitation in that activity and zero otherwise. We then sum the values to construct the ADL limitation score. Similarly, we construct the IADL limitation score, similar to Lawton and Brody's IADL index (Lawton and Brody, 1969), using four IADL-related questions (ability to shop, prepare meals, manage money, and take medications). In both cases, higher scores indicate more limitations. The ADL/IADL scores are similar to the Basel Index and the Lawton and Brody IADL Index, the leading indices in their respective fields, respectively. Still, they are not identical to them because of the limited number of questions included in the certification questions and differences in wording.

**The extent of Vertical Integration** We describe the extent of vertical integration of CMs using

the 2014 cross-section data. There are 553 CMs, and 76.7% vertically integrate at least one of the three services. Thus, vertical integration is quite common in Japanese at-home LTC services. The share of CMs who integrate one, two, and three service types are 44.7%, 22.1%, and 0.2%, respectively. Among the three services, CMs integrate homecare (44.7%) most often, followed by daycare (38.5%) and rehabilitation (6.2%).

# 4. Demand Inducement Effect of Vertical Integration

## 4.1. Does VI affect the service included in a care plan? (Service choice)

**Empirical Model** In our empirical analysis, we first examine whether vertical integration affects the type of service CMs include in a care plan. Specifically, we ask, do vertically integrated CMs create a care plan that includes the type of service provided by co-owned downstream firms? If so, this may indicate an agency problem because the recipient wants services most appropriate for them, not those that the CM's co-owned firm provides. We estimate the following linear model to assess this relationship:

$$Y_{ims} = \beta_0 + VI_{ms}\beta_s + \boldsymbol{X_i}\boldsymbol{\beta_i} + \boldsymbol{Z_m}\boldsymbol{\beta_m} + \boldsymbol{D_i}\boldsymbol{\beta_D} + \boldsymbol{A_t} + \boldsymbol{B_W} + \varepsilon_{ims}\,, \qquad (1)$$

where $Y_{ims}$ is the dummy variable that takes the value of one if user *i*'s CM *m* includes a service *s* (e.g., daycare) in the care plan at least once during the 12 months after the baseline certification and zero otherwise. $VI_{ms}$ is a vertical integration dummy variable that takes the value of one if CM *m* has service *s* (e.g., daycare) under the same ownership and zero otherwise.[20] $\beta_s$ is the coefficient of our main interest. $\boldsymbol{X_i}$ is a vector of user *i*'s attributes, including dummy variables for the care level, care minutes, age, age squared, gender, coinsurance rate, and dummy variables for the frequency of certification *i* received before the baseline certification. $\boldsymbol{Z_m}$ are a vector of CM *m*'s attributes, such as the number of CMs in the care management firm, ownership status (for-profit versus not-for-profit), and the number of recipients of the care management firm. We also include $\boldsymbol{D_i}$, the distance from *i* to the closest daycare and rehabilitation care providers (two variables) to control for access to care.[21] $\boldsymbol{A_t}$ is the year dummy variable. Although we examine a small area (an urban municipality), to further control for within-municipality geographical variations, we include six ward fixed effects, $\boldsymbol{B_W}$, which controls for resident characteristics such as average income; a ward is an administrative unit within a municipality. $\varepsilon_{ims}$ is an error term, which we cluster by CM *m*. We separately estimate this regression for each service. Note that this specification examines whether a specific service is included in a care plan and does not examine whether the CM refers the

[20] To be precise, we observe a care management office, not an individual care manager.
[21] We do not include distance to the closet home care provider because home care providers visit the recipient's residence, and therefore, distance to home care providers would not affect access to care for the recipient.

recipient to a co-owned provider.

We also estimate a variant of the model (1) to study "service substitution" by CMs. For example, suppose a CM has a home care service under the same ownership. Then, the CM may substitute home care for other services (e.g., daycare) if recipients are budget- or time-constrained. If true, this may also indicate an agency problem because vertical integration prevents the agent (CM) from including optimal services for the client (recipient).

To examine this possibility, we include additional vertical integration dummy variables for other services on the right-hand side and estimate regression (2).

$$Y_{ims} = \beta_0 + \sum_k VI_{mk}\beta_k + \boldsymbol{X_i\beta_i} + \boldsymbol{Z_m\beta_m} + \boldsymbol{D_i\beta_D} + \boldsymbol{A_t} + \boldsymbol{B_W} + \varepsilon_{ims}\,, \qquad (2)$$

where the vertical integration dummy variable, $VI_{mk}$, takes the value of one if CM $m$ and service $k$ (daycare, homecare, and rehabilitation) are under the same ownership and zero otherwise. A negative $\beta_k$ ($k \neq s$) indicates that the CM substitutes away from $s$ if the CM co-owns service $k$. In this regression, we focus on daycare, home care, and rehabilitation, the three popular services that recipients regularly use, usually two or three times weekly. If there is a service substitution, it would be between these three services.

**Identification** Our primary concern for identification is the recipients' selection of a CM. For example, recipients who prefer a particular service, say rehabilitation service, may choose a CM who has a rehabilitation service under the same ownership. Then, unobserved user preferences correlated with the vertical integration dummy variable affect the service included in the care plan, creating an endogeneity problem.

We address this endogeneity issue using differential distance (DD) instruments. Specifically, we calculate the distance from the user's residence to the closest vertically integrated and not-integrated CM for each service, take the difference, and use it as an instrument for vertical integration. We expect the shorter the relative distance to an integrated CM, the more likely the user will choose an integrated CM. The instrument is valid if this relationship holds and individuals do not choose a residence based on the distance to a CM. For robustness, we also experiment with two nonlinear versions of the DD instrument. The square-root version takes the square root of the differential distance and then takes the difference. The log version takes the logarithm of the distance (plus one) to the closest vertically integrated and not-integrated CM and takes the difference. For equation (2), we treat all three vertical integration dummy variables as endogenous and include three service-specific DD instruments. As shown below, the instruments also work in this case as CMs do not always integrate all services; thus, the instruments are not collinear. As discussed in the Introduction, previous studies on healthcare services have used the DD instrument, and they also work in our case, as

shown below.

**Descriptive Statistics** Table 1 shows how CMs include vertically integrated and not-integrated services in care plans. For example, the first row shows to what extent CMs who have daycare services under the same ownership include daycare (59.3%), rehabilitation (14.6%), and home care (40.1%) in their care plans. The Table shows a few clear patterns. First, CMs include each service in a care plan more often when the CM and a downstream service are co-owned. For example, the first two rows for daycare (column 1) show that CMs who integrate daycare service are 11.8 percentage points (ppts hereafter) (59.3% versus 47.5%) more likely to include daycare in a care plan than those who don't. We observe the same pattern for rehabilitation (23.8 ppts difference) and home care (11.8 ppts difference). Second, CMs who vertically integrate a service (e.g., daycare) are less likely to include other services in their care plans than CMs who do not integrate the service. For example, the third and fourth rows for daycare (column 1) show that CMs who vertically integrate rehabilitation are less likely to include daycare by 16.2 ppts (39.4% versus 55.6%) than those who don't. This pattern is consistent with vertically integrated CMs substituting away from non-vertically integrated services. Table 2 summarizes other characteristics of recipients and their CMs included in the "inducement effect" analyses.

**Instrument Validity** Figure 1 shows the distribution of the linear version of the DD instrument for the three services. They show that the DD instrument is distributed roughly evenly between -2 and 2 km for daycare and home care, while it is mainly positive for rehabilitation. The difference reflects that nearly half of CMs have daycare and home care services under the same ownership, while fewer CMs do so for rehabilitation. We also note many cases where the differential distance is zero. This typically happens when the closest CM with and without a co-owned service is in the same 7-digit postal code as the recipient. To check the robustness, we also estimate models excluding these cases, finding similar results.

Figure 2 plots the relationship between the DD instrument (linear version) and the vertical integration status of the CM. This is the first-stage relationship. All the figures show a clear negative relationship, which indicates that recipients are less likely to choose a vertically integrated CM if it is located far away relative to non-vertically integrated CMs, as expected.

Although the second assumption for the validity of an instrument that individuals do not choose residential locations based on the differential distance to VI and not-VI CMs cannot be directly tested, as reported in Table 3, user attributes below and above the median of the differential distance are generally balanced, supporting the assumption. Although a few variables are statistically significant due to large observations, the absolute differences are small. Nonetheless, we include these variables as regressors to control for the differences in attributes.

**Estimation Results** Table 4 reports results from equation (1). Panels A, B, and C report results for daycare, rehabilitation, and home care, respectively. Each panel reports OLS, first-stage, and second-stage results, respectively. We show three alternative specifications for the DD instrument: the linear, square-root, and log versions. For brevity, we only report the coefficients on the vertical integration dummy variable and the DD instrument.

We first look at the first-stage results. The estimated coefficients on the DD instrument are all negative, as expected, indicating that integrated CMs are less likely to be selected if they are located relatively far from non-integrated CMs. Among the three versions of the DD instruments, the square-root version tends to have higher t-statistics than the other two. For this reason, we mainly use the square-root version throughout the paper.

The coefficients on the vertical integration dummy variable (the second stage results) are all positive and statistically significant, except for daycare with the linear DD instrument, indicating that vertical integration generally increases the probability of including a co-owned service in the care plan. If we focus on the results with the square-root version of the DD instrument (column 5), where the first-stage results are stronger, the probability increases from 29% to 60%, depending on the service. The positive effect of vertical integration is consistent with the descriptive statistics discussed above. Appendix Table A** (to be added) reports results when excluding observations where both VI and not-VI CMs are in the same 7-digit ZIP code as the recipient, and thus the differential distance is zero. The results change little with this sample, providing further confidence in our main results.

Table 5a reports results from equation (2), which includes three VI dummy variables as regressors. The name of the service in each column corresponds to the dependent variable of the regression (whether the service is included in a care plan), while the service names in rows correspond to the independent variables (whether the CM vertically integrates daycare, rehabilitation, and home care). As before, we report results using the square-root version of the DD instrument. We first look at the results for daycare (columns 1 and 2). The OLS results indicate that daycare service is more likely to be included in a care plan when CMs integrate a daycare service. Interestingly, CMs are less likely to include daycare in a care plan when CMs integrate other services, i.e., rehabilitation and home care. This result suggests CMs substitute services in favor of integrated services.

The same pattern holds for the 2SLS results (column 2), although the substitution effect becomes insignificant. The results are similar for rehabilitation (columns 3 and 4) and home care (columns 5 and 6). The estimated impact of vertical integration on care plan choice is large; co-ownership increases the probability of including daycare in a care plan by as much as 29%. The corresponding numbers for rehabilitation and home care are 49% and 41%. These results are at least weekly significant. Turning to the substitution effects, we find that all the coefficients are negative, and the point estimates range from -7% to -48%, although most of them are not statistically significant. The F-statistics are larger than 24, indicating that the

instrumental variables are relevant.

**Heterogeneity by dementia level** Table 5b reports results when we split the sample by the level of dementia dependency and estimate equation (2). Panel A (B) reports results for recipients with higher (lower) dementia defined by the level of dementia dependency greater than or equal to 2 (lower than 2). As discussed in Section 2.2, Level 0 recipients have no dementia, and Level 1 recipients have some form of dementia, but they are largely independent in daily life, both in the home and socially. We split the sample this way so that the two groups have approximately equal samples.

Focusing on the 2SLS results, we find that the results for the higher dementia group (Panel A) are qualitatively the same as the average effect (Table 5a), and the point estimates are somewhat larger than the average effect. Specifically, we continue to find that vertical integration induces integrated CMs to include co-owned services in a care plan. We also find suggestive evidence for service substitution – integrated CMs substitute not-co-owned services with co-owned services – although the estimates are significant only in a few cases. In contrast, Panel B results for low dementia recipients indicate that the demand inducement effect of vertical integration is weaker. For example, CMs who vertically integrate daycare and rehabilitation services are no longer more likely to include respective services in a care plan. In both cases, the first-stage F-statistics are larger than 10. These results suggest that the demand inducement effect found in Table 5 is primarily driven by recipients with a higher level of dementia.

An important factor that may affect CMs' decisions is the presence of family members. We would expect agency problems, if any, would be alleviated if recipients with cognitive limitations had a family member who could monitor the CM's decisions. Unfortunately, we do not have information about family members in the data, such as whether the recipient lives with a child. However, studies have shown that recipients with higher cognitive limitations are more likely to live with family members (e.g., Kobayashi et al., 2011). Then, our estimates on the impact of vertical integration for recipients with a higher level of dementia are likely to be conservative because the estimates would be biased toward zero to the extent family members mitigate the agency problem.

To check robustness, we also estimate the same model by dividing the sample by the IADL score (see Section 3 for details), an alternative measure of cognitive limitations. Specifically, we consider those with an IADL score of four (approximately 63% of the sample) as having higher IADL limitations and those with less than four (approximately 37%) as having lower IADL limitations. Appendix Table A1 reports the results. We find that the results are similar to those using dementia dependency as the measure of cognitive limitations. In particular, the 2SLS results indicate that the impact of vertical integration on including a co-owned service type in a care plan is more pronounced when recipients have more cognitive limitations

measured by the IADL score. It is reassuring that the results change little, regardless of the specific measure of cognitive limitations.

To summarize, we find evidence that vertically integrated CMs are more likely to include co-owned services in their care plans, and the impact is stronger for recipients with higher cognitive limitations. There is also suggestive evidence that vertically integrated CMs tend not to include not-co-owned services in a care plan, although the estimates are largely insignificant.

### 4.2. How much does VI affect the volume of service? (Service Volume)

**Empirical Model and Data** In the previous section, we estimated the impact of vertical integration at the extensive margin, i.e., whether vertically integrated CMs include a co-owned service in a care plan. This section extends the analysis and examines the impact of vertical integration on service volume, i.e., the intensive margin. Our empirical model is essentially the same as equation (2) above. The only difference is that we replace the dependent variable, $Y_{ims}$, by recipient *i*'s average monthly LTC expenditure for service *s* (such as adult daycare) for twelve months following the baseline certification. Additionally, we estimate a model that uses the sum of the three service expenditures as the dependent variable. The sum of expenditures will increase if vertically integrated CMs induce demand for vertically integrated services on top of all other service expenditures. Alternatively, it will not increase if CMs substitute vertically integrated services with not-vertically integrated services. We use the same data as before. Please see Table 2 for the summary statistics on average monthly expenditure.

**Estimation Results** Table 6 reports estimation results for service volume. As before, we use the square-root version of the DD instrument for identification. The first two columns report results for daycare services. The OLS results in column 1 suggest that daycare spending is higher if CMs vertically integrate daycare services. At the same time, daycare spending is lower if CMs integrate rehabilitation or home care services, which indicates service substitution. Column 2 shows that these results largely hold even when we use instrumental variables. The 2SLS estimates in column 2 suggest that daycare service expenditure increases by 4,946 points per month (approximately 49,000 JPY or 445 USD) if CMs co-own daycare services but decreases by 4,883 points per month (approximately 48,800 JPY or 444 USD) if CMs integrate home care. These estimates are statistically significant. Again, the latter result indicates that vertically integrated CMs substitute not-integrated services with integrated services. The F-statistics reported at the bottom of the table are larger than 24, indicating that the instrumental variables are relevant.

The estimated impacts are large, considering that the average monthly spending is 3,000 points for daycare, as reported in Table 2. However, the large impact may not be surprising because, as we saw in Section 4.1, there is a large extensive margin effect; vertically integrated CMs are likely to include co-owned services in a care plan but are less likely to do so for not-

co-owned services. These results also indicate that the volume of a service (e.g., daycare) drastically changes depending on what service the CM vertically integrates.

We find the same pattern – CMs increase service volume for the service they integrate with while reducing service volume for not-integrated services – holds for the other two services, albeit somewhat less. For example, the results in column 6 for home care indicate that CMs that vertically integrate home care significantly increase home care services, while those that integrate day care and rehabilitation decrease home care services, although the latter two are not statistically significant.

To further examine the impact of vertical integration on total expenditure, Table 6, column 8 reports the 2SLS results when we use the sum of the three service expenditures as the dependent variable. The 2SLS results in Panel B indicate that vertical integration of any services does not significantly affect the sum of expenditures. This result further supports that vertically integrated CMs are likely to substitute not-integrated services with integrated services without substantially affecting the total LTC expenditures.

**Heterogeneity by dementia level** Table 7 reports results when we split the sample by the level of dementia dependency and estimate the same model as the above. The 2SLS results in Panel A for higher dementia recipients are qualitatively the same as those in the full sample reported in Table 6. Specifically, LTC expenditure of a service (e.g., daycare) increases significantly when CMs co-own the service (e.g., daycare), and there is suggestive evidence for service substitution that integrated CMs tend to reduce the service volume of not-co-owned services. Interestingly, the results in Panel B for lower dementia recipients are in stark contrast to those in Panel A. In particular, none of the 2SLS results are statistically significant, indicating that there is no evidence that vertically integrated CMs behave differently from not-integrated CMs for lower dementia recipients.

Appendix Table A2 reports results when we use the alternative measure of cognitive limitations, the IADL score. We continue to find that integrated CMs are more likely to include co-owned service types in a care plan when recipients have higher cognitive limitations, although our estimates for the low IADL group are less reliable with the first-stage F-statistics as low as 8.5.

In sum, the volume effect analysis (intensive margin) indicates that vertical integration increases the volume of service CMs co-own and reduces non-co-owned volumes. These effects cancel each other without increasing the total LTC expenditure. These results suggest that the primary impact of vertical integration is substituting not-integrated services with integrated services, which is different from the traditional demand inducement effect of vertical integration that increases the volume of the same service. The fact that vertical integration distorts CMs' choice of service and service volume is concerning. However, it is also possible that vertical integration improves efficiency and benefits the recipient. We now analyze whether vertical

integration improves efficiency regarding recipients' care needs and health outcomes.

## 5. Efficiency Effect of Vertical Integration

**Empirical Model** We now turn to the efficiency effects of vertical integration. We examine whether vertical integration between CMs and providers slows the i) progression of care needs and ii) deterioration of health outcomes. By eliminating organizational boundaries, vertical integration may improve care coordination, which, in turn, may delay the progression of care needs and the deterioration of health outcomes. If so, it could also help reduce future LTC expenditures, as LTC expenditures would increase with care needs. We examine these possibilities by estimating the following linear model:

$$\Delta Y_{imt} = \beta_0 + VI_m^{any}\beta_{VI} + \boldsymbol{X_i\beta_i} + \boldsymbol{Z_m\beta_m} + \boldsymbol{D_i\beta_D} + \boldsymbol{A_t} + \boldsymbol{\varepsilon_{imt}}\,, \qquad (3)$$

where $\Delta Y_{imt}$ is recipient *i*'s change in care needs or health outcomes 14 (or 26) months after *i*'s baseline certification at *t*, where $\Delta Y_{imt}$=$Y_{im,t+14} - Y_{imt}$ (for the 14 months data). The subscript *m* stands for *i*'s CM. We measure the change in the outcomes at the individual level as they are not separately available for each service. We discuss the dependent variables in detail below. We study the outcomes 14 and 26 months after the baseline certification because most recipients receive at least one (two) recertification(s) within 13 (25) months. Thus, the dependent variables, such as care needs and health outcomes, are updated at least once by then. $VI_m^{any}$ is a dummy variable for vertical integration, which takes the value of one if recipient *i*'s CM *m* has any of the three services (daycare, home care, rehabilitation service) under the same ownership and zero otherwise. $\beta_{VI}$ is the coefficient of our main interest. We also estimate a model that separately includes the three vertical integration statuses; for this model, we replace $VI_m^{any}\beta_{VI}$ by $\sum_s VI_{ms}\beta_{VI_s}$, where *s* stands for service. The other regressors are the same as before. Again, we cluster the error term by CM, *m*.

**Dependent variables** We examine two types of dependent variables: i) care needs and ii) health outcomes. Although these two outcomes are closely related, they are different as the former relates to the burden of care, and the latter captures the degree of frailty or health status of the recipient. For care needs, we use required care minutes evaluated by municipal investigators during certification as the primary dependent variable. Additionally, we create a dummy variable, *Worsen*, which takes the value of one if recipient *i*'s care level worsens relative to baseline and zero otherwise.[22] For health outcomes, we use three variables: ADL/IADL

[22] For this analysis, we exclude recipients with care level=5 at the baseline as their care level cannot worsen from 5.

limitation score and dementia dependency.

**Identification** Our primary endogeneity concern is the selection of CMs by recipients. As discussed in Section 4.1, recipients may have unobservable preferences for vertically integrated CMs. If they directly affect the outcomes of our interest, the estimated coefficient on the vertical integration status will be biased. We address this issue using again the DD instrument. This time, we modify the DD instrument because our vertical integration dummy, $VI_m^{any}$, takes the value of one if $m$ integrates any of the three services. We calculate the differential distance between the closest CM that vertically integrates any of the three services and does not integrate any and use it as an instrument. Nonetheless, as before, we expect that the instrument negatively relates to $VI_m^{any}$; vertically integrated CMs are less likely to be chosen if they are located relatively far away than not-vertically integrated CMs. For the analysis that separately includes the three vertical integrations status, we use the DD instruments defined separately for each service as in Section 4.1.

Another identification issue is that other unobserved CM attributes, which may independently affect the outcomes, may correlate with the vertical integration status. For example, if vertically integrated CMs are associated with higher-quality providers, it may be difficult to distinguish the impact of vertical integration from unobserved provider quality. We address this issue by following Card et al. (2023). Specifically, we control CM quality by including a high (versus low) quality dummy for CM in addition to the vertical integration status. Specifically, we estimate the following equation (4):

$$\Delta Y_{imt} = \beta_0 + VI_m^{any}\beta_{VI} + Q_{-i,m}\beta_q + \boldsymbol{X_i\beta_i} + \boldsymbol{Z_m\beta_m} + \boldsymbol{D_i\beta_D} + \boldsymbol{A_t} + \boldsymbol{\varepsilon_{imt}}\,, \qquad (4)$$

where $Q_{-i,m}$ is a dummy variable that takes the value of one if recipient $i$'s CM $m$ is a high-quality CM. To construct the quality indicator, we first calculate the care-minute change for each recipient and take the average by CM, excluding recipient $i$. We then divide CMs between high- versus low-quality types by the median of the quality indicator.[23] In our context, there are potentially multiple quality indicators, including care minutes, ADL, IADL, and dementia dependency. Among these, we use care minutes as the quality indicator, the primary measure capturing the level of care needs. It is also suitable because of the larger variance, partly owing to the measure's granularity (0.1minute). We calculate a CM's quality by calculating each recipient's change in care minutes during the first year after baseline certification, take the average for each CM, and divide CMs between high- versus low-quality types by the median. Then, recognizing that the quality dummy is also endogenous, we construct an additional DD instrument for the quality dummy in the same way as before by measuring the relative distance

[23] We construct this index by focusing on CMs who have at least 30 recipients in our data. The results are robust to a lower threshold of 10 recipients, which are available from the authors upon request.

from a recipient's location to the closest high-- versus low-quality type CMs.

**Descriptive Statistics** Table 8 reports summary statistics. We focus on the recipients who have at least one of the three services in the care plan and do not use other services. The Table shows that the mean of $VI_m^{any}$ is 0.785, indicating that 79% of CMs vertically integrate at least one of the three services. CMs integrate daycare and home care services more frequently (50.6% and 52.7%, respectively) than rehabilitation (14.2%).

Figure 3 shows the relationship between the differential distance (linear version) and the probability that a recipient chooses a CM who vertically integrates any of the three services. This figure shows that the differential distance is negatively correlated with the choice of vertically integrated CM, as expected, validating the relevance of the DD instrument. Table 9 reports that user attributes below and above the median of differential distance are generally balanced, supporting the assumption. Although a few variables show statistical significance due to large observations, the absolute differences are small.

Figure 4 looks at the relationship between the differential distance (linear version) and the probability that recipients choose a high-quality CM. Whether we define quality by care minutes (Panel A) or ADL limitation (Panel B), there is a negative relationship between the two, indicating that high-quality CMs are more likely to be chosen if they are more closely located than low-quality CMs, as expected.

**Estimation Results** Table 10 reports estimation results for efficiency effects using outcomes data 14 months after the baseline certification. As before, we use the square-root version of the instrument, which had higher t-statistics in Table 4. For each outcome, the first column presents the first-stage result that considers only the vertical integration status as endogenous. The second column presents the second-stage result based on the same assumption. The third column reports the second-stage result when we include the high-quality CM dummy as an additional endogenous variable.

First, the first-stage result indicates that, for both care needs and health outcomes, the t-statistics for the DD instrument are almost four in all cases, indicating that the instruments have power. Turning to the second column of each outcome, we find that the estimated coefficients are none of the coefficients are significant at the conventional level. Recall that a larger value of the dependent variable corresponds to a higher level of frailty and dementia. These results indicate that there is no evidence that vertical integration improves efficiency in terms of the recipient's care needs and health outcomes. The third column of each outcome includes CM quality as an additional endogenous variable. The coefficients on the vertical integration dummy variable remain insignificant even after controlling for another potentially endogenous CM quality. We note that the coefficients on the high-quality CM dummy variable are negative in all cases and sometimes statistically significant, indicating that high-quality CM helps slow

the progression of care needs and deterioration of health outcomes, as expected. The F-statistics reported at the bottom of the table are larger than at least 18, indicating that the instruments are still relevant even after adding the additional instrument.

Table 11 shows the medium-run results after 26 months since the baseline certification. The results are similar to the short-run results reported in Table 10; the coefficients on the vertical integration dummy variable (the second column of each outcome) are not negative and statistically significant in all cases, indicating that there is no evidence that vertical integration improves the efficiency measures we examine. If anything, the positive and significant coefficient on ADL suggests that vertical integration worsens the frailty measure, although the result is not robust across different models.

Table 12 reports the results that separately include the three vertical integration dummy variables as the independent variables. This specification is potentially useful in identifying the differential impact of vertical integration across the three services. We find that none of the estimated coefficients from the 2SLS model is statistically significant, supporting the results reported in Table 11. The F-statistics, reported at the bottom of the table, are approximately 10 across the models, indicating that the service-specific DD instruments are still relevant.[24] The medium-run results are similar, and we report them in Table 13.

**Heterogeneity by dementia level** Table 14 reports results when we split the sample by the level of dementia dependency and estimate the same model as the above. The results in Panel A for higher dementia recipients are similar to those in the full sample reported in Table 10. First, the first-stage result indicates that, for both care needs and health outcomes, the t-statistics for the DD instrument are larger than four in all cases, indicating that the instruments have power. Turning to the second column of each outcome, we find that the estimated coefficients are positive, but none are statistically significant. Again, we find no evidence that vertical integration improves efficiency regarding the recipient's care needs and health outcomes. The third column of each outcome includes CM quality as an additional endogenous variable. The coefficients on the vertical integration dummy variable remain insignificant even after controlling for CM quality. Panel B reports results for lower dementia recipients. Although estimated coefficients show a similar pattern, they are less reliable as the instruments are not strong enough, as indicated by the small F-statistics less than 10.[25]

In summary, our results do not provide evidence that vertical integration improves efficiency regarding the recipient's care needs and health outcomes.

---

[24] For completeness, we also estimate the models that include the high-quality CM dummy variable, another endogenous regressor. Unfortunately, our instruments are no longer strong enough with the F-statistics smaller than 10, and the results are not reported.

[25] We also report the efficiency effects after 26 months of baseline certification reported in Table 15. The results are similar; none of the estimated coefficients are negative and significantly different from zero.

# 6. Additional Analysis and Robustness Check

## 6.1. Attrition

In Appendix D, we consider two attrition issues. As discussed in Section 3, our inducement and efficiency effects analyses include individuals certified as of the 12th month and 14$^{th}$ month since baseline certification, respectively. Then, a sample selection problem may be a concern if the attrition rate, e.g., not certified as of the 12$^{th}$ month, differs by the vertical integration status. To examine differential attrition, we construct the attrition dummy variable that equals one if the recipient has a baseline certification but is not included in the respective estimation sample and zero otherwise. Then, we estimate equations (2) and (3) using the attrition dummy variable as the dependent variable. Appendix Table D1 reports estimation results. The results in Panel A for the inducement effect show that there is no differential attrition by the vertical integration status. Panel B results for the efficiency effect are similar; there is no evidence of differential attrition at the conventional significance level. Therefore, sample selection problems due to attrition are not a major issue in our samples.

## 6.2. Robustness Checks

To be added

# 7. Conclusion

Do vertically integrated agents work in the best interest of their clients? This paper examined this issue in the context of CMs coordinating LTC services for frail older people who often have cognitive limitations. Many alternative at-home or non-institutional LTC services are available, and choosing an appropriate one is important but challenging for frail older people. Therefore, care management is a crucial function in all aging societies, although who plays the role differs by the specific institutional setup. A concern for vertical integration is that integrated agents may recommend a service to benefit co-owned downstream firms even if it is inappropriate for recipients. The concern is even stronger if recipients have cognitive limitations and have difficulty evaluating agents' recommendations. On the other hand, vertical integration may benefit recipients by promoting better coordination or information transmission within the boundaries of the firm. This, in turn, may slow the progression of care needs and the deterioration of health outcomes. No previous studies have examined the tradeoff despite the need for efficient care coordination in LTC services, an important emerging sector in all aging societies. Moreover, little is known about agency conflicts when clients have cognitive limitations.

Using data from Japan, we study the impact of vertical integration of care management on the demand inducement and efficiency effects. In contrast to previous studies that show volume increase as primary evidence of demand inducement, we examine two types of demand

inducement effects: the type of service CMs include in a care plan and the volume of these services CMs recommend. Studying the service CMs choose is important because selecting an appropriate service is crucial for CMs who represent frail older people who may have difficulty selecting appropriate services.

We find that vertical integration of care management affects the type of service and the volume of services recipients receive. In particular, vertically integrated CMs are more likely to include co-owned downstream services in a care plan and exclude not-co-owned service types. We find the same pattern for the volume of services. Moreover, the demand inducement effect is more prominent for recipients with higher cognitive limitations. Interestingly, LTC expenditure in total does not significantly differ with and without vertical integration. These results indicate that the primary effect of integrating CMs is to substitute not-co-owned services with co-owned services.

The fact that vertically integrated CMs distort care plans in favor of co-owned downstream services suggests that vertical integration costs to recipients, although LTC spending in total, do not appear to increase. However, vertical integration may still be acceptable or even recommended if its efficiency gains outweigh its costs. We study this possibility empirically and find no evidence that vertical integration improves efficiency by slowing the progression of care needs or deterioration of health outcomes relative to not-integrated CMs. Thus, our results indicate that vertical integration of care management leads to an agency problem by distorting the services recipients receive without clear compensating benefits to recipients, and the impact is larger when clients have more cognitive limitations.

We believe this paper enhances the understanding of the principal-agent relationship, particularly when principals face cognitive limitations. However, the current research has its limitations. For example, agency problems might be alleviated if a principal has a guardian or family member who can monitor the agent's behavior. Investigating whether and to what extent this is true poses an important empirical question that could help improve organizational structures in aging societies. Nevertheless, this critical question is left for future research.

## References


Afendulis, C. C., & Kessler, D. P. (2007). Tradeoffs from integrating diagnosis and treatment in markets for health care. American Economic Review, 97(3), 1013-1020.

Baker, L. C., Bundorf, M. K., & Kessler, D. P. (2014). Vertical integration: hospital ownership of physician practices is associated with higher prices and spending. Health Affairs, 33(5), 756-763.

Baker, L. C., Bundorf, M. K., & Kessler, D. P. (2016). The effect of hospital/physician integration on hospital choice. Journal of Health Economics, 50, 1-8.

Bakx, P., Wouterse, B., Van Doorslaer, E., & Wong, A. (2020). Better off at home? Effects of nursing home eligibility on costs, hospitalizations and survival. Journal of Health Economics, 73, 102354.

Barro, J. R., Huckman, R. S., & Kessler, D. P. (2006). The effects of cardiac specialty hospitals

on the cost and quality of medical care. Journal of health economics, 25(4), 702-721.
Beckert, W., Christensen, M., & Collyer, K. (2012). Choice of NHS-Funded hospital services in England. The Economic Journal, 122(560), 400-417.
Banerjee, S., B. P. Carlin and A. E. Gelfand (2004) Hierarchical Modeling and Analysis for Spatial Data, Chapman and Hall / CRC Press.
Campbell, J. C., Ikegami. N. and Gibson. M. J. (2010). Lessons from public long-term care insurance in Germany and Japan. Health Affairs, 29(1), 87-95.
Card, D., Fenizia, A., & Silver, D. (2023). The health impacts of hospital delivery practices. American Economic Journal: Economic Policy, 15(2), 42-81.
Chandra, Amitabh, Courtney Coile, and Corina Mommaerts. 2023. "What Can Economics Say about Alzheimer's Disease?" Journal of Economic Literature, 61 (2): 428–70.
Cutler, D. M., Dafny, L., Grabowski, D. C., Lee, S., & Ody, C. (2020). Vertical integration of healthcare providers increases self-referrals and can reduce downstream competition: the case of hospital-owned skilled nursing facilities. National Bureau of Economic Research WP w28305.
De Meijer, C., Bakx, P., Van Doorslaer, E., & Koopmanschap, M. (2015). Explaining declining rates of institutional LTC use in the Netherlands: a decomposition approach. Health Economics, 24, 18-31.
Fu, R., Iizuka, T., & Noguchi, H. (2023). Long-Term Care in Japan. In K. McGarry & J. Gruber (Eds.). Long-Term Care Around the World. University of Chicago Press, forthcoming.
Hisano, S. (2009). The relationship between Revised Hasegawa Dementia Scale (HDS-R), Mini-Mental State Examination (MMSE), Bed-Fast Scale, and Dementia Scale. Jpn J Geriatr Psychiatry, 20, 883, (in Japanese).
Ho, K., & Pakes, A. (2014). Physician payment reform and hospital referrals. American Economic Review, 104(5), 200-205.
Hsu H.C and Chen, C. F. (2019), "LTC 2.0: The 2017 reform of home- and community-based long-term care in Taiwan," Health Policy, 123; 912-916.
IHEP (2020) “ケアマネジメントの公正中立性を確保するための取り組みや質に関する
Iizuka, T. (2007). Experts' agency problems: evidence from the prescription drug market in Japan. The Rand journal of economics, 38(3), 844-862.
Iizuka, T. (2012). Physician agency and adoption of generic pharmaceuticals. American Economic Review, 102(6), 2826-2858.
Johri, M., Beland, F., and Bergman, H. (2003). International experiments in integrated care for the elderly: a synthesis of the evidence. International Journal of Geriatric Psychiatry, 18(3), 222-235.
Kim, H., and Kwon, S. (2021). A decade of public long-term care insurance in South Korea: Policy lessons for aging countries. Health Policy, 125(1), 22-26.
Kishida, K. (2015). Supplier-induced demand for in-home elderly care (in Japanese), Iryo Keizai Kenkyu, 27(2),117-134
Kobayashi, E., Fujiwara, Y., Fukaya, T., Nishi, M., Saito, M. and Shinkai, S. (2011) Social support availability and psychological well-being among the socially isolated elderly: Differences by living arrangement and gender (in Japanese), Nihon Koushuu Eisei Zasshi, 58(6), 446-456
Koch, T. G., Wendling, B. W., & Wilson, N. E. (2017). How vertical integration affects the quantity and cost of care for Medicare beneficiaries. Journal of health economics, 52, 19-32.
Koch, T. G., Wendling, B. W., & Wilson, N. E. (2021). The Effects of Physician and Hospital Integration on Medicare Beneficiaries' Health Outcomes. Review of Economics and Statistics, 103(4), 725-739.
Konetzka, R. T., Stearns, S. C., & Park, J. (2008). The staffing–outcomes relationship in nursing homes. Health services research, 43(3), 1025-1042.
Konetzka, R. T., Stuart, E. A., & Werner, R. M. (2018). The effect of integration of hospitals and post-acute care providers on Medicare payment and patient outcomes. Journal of health

economics, 61, 244-258.
Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. The Gerontologist, 9(3_Part_1), 179–186.
Lin, H., McCarthy, I. M., Richards, M. R., & Whaley, C. (2022). Owning the agent: hospital influence on physician behaviors. National Bureau of Economic Research WP w28859.
Mahoney, F. I., & Barthel, D. W. (1965). Functional evaluation: the Barthel Index: a simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill. Maryland State Medical Journal.
McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior, Chapter in Frontiers in Econometrics (P. Zarembka, ed.), Academic Press New York, 1973.
McClellan, M., McNeil, B. J., & Newhouse, J. P. (1994). Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality?: analysis using instrumental variables. Jama, 272(11), 859-866.
Ministry of Health, Labour and Welfare. (2004). *Survey on Care Managers.* In Japanese. https://www.mhlw.go.jp/shingi/2004/02/s0223-8i3.html (accessed Dec 6, 2023).
Mitsubishi Sougou Kenkyuujo (2020) *Survey on Self Care Plans*. In Japanese. https://pubpjt.mri.co.jp/pjt_related/roujinhoken/jql43u00000000s3-att/R1_053_4_manual.pdf (accessed Dec 6, 2023).
Nakamura, S., Capps, C., & Dranove, D. (2007). Patient admission patterns and acquisitions of "feeder" hospitals. Journal of Economics & Management Strategy, 16(4), 995-1030.
Rahman, M., Grabowski, D. C., Gozalo, P. L., Thomas, K. S., & Mor, V. (2014). Are dual eligibles admitted to poorer quality skilled nursing facilities?. Health services research, 49(3), 798-817.
Rahman, M., Norton, E. C., & Grabowski, D. C. (2016). Do hospital-owned skilled nursing facilities provide better post-acute care quality?. Journal of Health Economics, 50, 36-46.
Sugawara, S. and Nakamura, J. (2016) Gatekeeper incentives and demand inducement: An empirical analysis of care managers in the Japanese Long-Term Care Insurance program, Journal of the Japanese and International Economies, 40, 1-16.
Tay, A. (2003). Assessing competition in hospital care markets: the importance of accounting for quality differentiation. RAND Journal of Economics, 786-814.
Tsutsui, T.and N. Muramatsu. 2005. "Care-needs certification in the long-term care insurance system of Japan." Journal of the American Geriatrics Society 53(3):522-527.
Wooldridge, J. M. (2014). Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables. Journal of Econometrics, 182(1), 226-234.
Wooldridge, J. M. (2015). Control function methods in applied econometrics. Journal of Human Resources, 50(2), 420-445.

# Tables

Table 1: Probability of including a service in a care plan by the CM's VI status

| | Fraction included in the care plan | | | |
|---|---|---|---|---|
| | Daycare | Rehabilitation | Home care | N |
| Daycare VI | 0.593 | 0.146 | 0.401 | 13,946 |
| Daycare non-VI | 0.475 | 0.204 | 0.456 | 14,582 |
| Rehabilitation VI | 0.394 | 0.380 | 0.316 | 4,089 |
| Rehabilitation non-VI | 0.556 | 0.142 | 0.448 | 24,439 |
| Home care VI | 0.542 | 0.148 | 0.486 | 14,768 |
| Home care non-VI | 0.522 | 0.205 | 0.368 | 13,760 |

*Notes*: This table shows how care managers include vertically integrated and not-integrated services in their care plans. For example, the first row shows the extent to which care managers with a daycare service under the same ownership include daycare, rehabilitation, and home care in their care plans. N represents the number of recipients.

Table 2: Summary statistics (inducement effects)

| | Mean | SD | Min | Max | N |
|---|---|---|---|---|---|
| Use of daycare | 0.533 | 0.499 | 0 | 1 | 28,528 |
| Use of rehabilitation | 0.176 | 0.381 | 0 | 1 | 28,528 |
| Use of home care | 0.429 | 0.495 | 0 | 1 | 28,528 |
| Spending on daycare (point) | 3,004 | 4,578 | 0 | 36,788 | 28,528 |
| Spending on rehabilitation (point) | 894 | 2,475 | 0 | 26,800 | 28,528 |
| Spending on home care (point) | 2,314 | 5,227 | 0 | 49,611 | 28,528 |
| Spending on three services (point) | 6,211 | 6,590 | 0 | 51,524 | 28,528 |
| VI daycare | 0.489 | 0.500 | 0 | 1 | 28,528 |
| VI rehabilitation | 0.143 | 0.350 | 0 | 1 | 28,528 |
| VI home care | 0.518 | 0.500 | 0 | 1 | 28,528 |
| DD daycare | -0.010 | 0.630 | -4.478 | 2.450 | 28,528 |
| DD rehabilitation | 1.035 | 0.846 | -3.196 | 4.371 | 28,528 |
| DD home care | 0.017 | 0.629 | -1.641 | 4.845 | 28,528 |
| Age | 81.061 | 7.536 | 65 | 105 | 28,528 |
| Female | 0.603 | 0.489 | 0 | 1 | 28,528 |
| Coins 0% | 0.092 | 0.289 | 0 | 1 | 28,528 |
| Coins 10% | 0.811 | 0.392 | 0 | 1 | 28,528 |
| Coins 20% | 0.091 | 0.287 | 0 | 1 | 28,528 |
| Coins 30% | 0.014 | 0.116 | 0 | 1 | 28,528 |
| Care level | 2.086 | 1.239 | 1 | 5 | 28,528 |
| Care minutes | 60.570 | 25.824 | 32 | 173.8 | 28,528 |
| ADL | 4.432 | 1.608 | 0 | 7 | 28,528 |
| IADL | 3.330 | 1.014 | 0 | 4 | 28,528 |
| Dementia | 2.667 | 0.974 | 1 | 6 | 28,519 |

*Notes*: This table reports the summary statistics of the service choice estimation samples. Spending represents average monthly spending on services. The unit of spending is points, and 1 point is approximately 10 yen. N represents the number of recipients.

Table 3: User attributes below and above the median of differential distance (service choice)

(A) Daycare

| | DD < median | DD >= median | Difference | p-value |
|---|---|---|---|---|
| Age | 80.934 | 81.107 | -0.173 | 0.086 |
| Female | 0.604 | 0.603 | 0.001 | 0.900 |
| Coins 0% | 0.092 | 0.092 | 0.000 | 0.972 |
| Coins 10% | 0.820 | 0.808 | 0.012 | 0.022 |
| Coins 20% | 0.083 | 0.094 | -0.010 | 0.009 |
| Coins 30% | 0.012 | 0.014 | -0.002 | 0.222 |
| Care level | 2.098 | 2.081 | 0.017 | 0.294 |
| Care minutes | 60.807 | 60.483 | 0.323 | 0.350 |
| ADL | 4.450 | 4.426 | 0.024 | 0.256 |
| IADL | 3.342 | 3.325 | 0.017 | 0.212 |
| Dementia | 2.673 | 2.664 | 0.009 | 0.514 |
| N | 7,595 | 20,933 | | |

(B) Rehabilitation

| | DD < median | DD >= median | Difference | p-value |
|---|---|---|---|---|
| Age | 81.073 | 81.048 | 0.025 | 0.777 |
| Female | 0.609 | 0.598 | 0.010 | 0.076 |
| Coins 0% | 0.096 | 0.087 | 0.009 | 0.008 |
| Coins 10% | 0.808 | 0.813 | -0.005 | 0.272 |
| Coins 20% | 0.088 | 0.094 | -0.006 | 0.092 |
| Coins 30% | 0.014 | 0.013 | 0.002 | 0.236 |
| Care level | 2.073 | 2.098 | -0.025 | 0.089 |
| Care minutes | 60.337 | 60.802 | -0.465 | 0.129 |
| ADL | 4.414 | 4.451 | -0.037 | 0.052 |
| IADL | 3.333 | 3.326 | 0.008 | 0.526 |
| Dementia | 2.652 | 2.681 | -0.029 | 0.013 |
| N | 14,257 | 14,271 | | |

(C) Home care

| | DD < median | DD >=median | Difference | p-value |
|---|---|---|---|---|
| Age | 81.151 | 81.025 | 0.126 | 0.203 |
| Female | 0.606 | 0.602 | 0.004 | 0.519 |
| Coins 0% | 0.082 | 0.096 | -0.014 | 0.000 |
| Coins 10% | 0.822 | 0.806 | 0.016 | 0.002 |
| Coins 20% | 0.090 | 0.091 | -0.001 | 0.747 |
| Coins 30% | 0.014 | 0.014 | 0.000 | 0.858 |
| Care level | 2.108 | 2.077 | 0.031 | 0.058 |
| Care minutes | 60.866 | 60.452 | 0.414 | 0.222 |
| ADL | 4.469 | 4.418 | 0.051 | 0.016 |
| IADL | 3.341 | 3.325 | 0.016 | 0.220 |
| Dementia | 2.680 | 2.661 | 0.019 | 0.147 |
| N | 8,098 | 20,430 | | |

*Notes*: This table reports the user attributes below and above the median of the differential distance for each service. We also report the p-value of the t-test for the difference between the means. N represents the number of recipients.

Table 4: Estimation results for "service choice"

| | | 2SLS | | | | | |
|---|---|---|---|---|---|---|---|
| | | Instrument = linear | | Instrument = square root | | Instrument = log | |
| | OLS | First stage | Second stage | First stage | Second stage | First stage | Second stage |
| **Panel A (Daycare)** | | | | | | | |
| VI Daycare | 0.119*** | | 0.744 | | 0.330** | | 0.404** |
| | (0.0146) | | (0.467) | | (0.132) | | (0.161) |
| DD | | -0.0389** | | -0.0645*** | | -0.0891*** | |
| | | (0.0186) | | (0.0151) | | (0.0232) | |
| N | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 |
| **Panel B (Rehabilitation)** | | | | | | | |
| | | 2SLS | | | | | |
| | | Instrument = linear | | Instrument = square root | | Instrument = log | |
| | OLS | First stage | Second stage | First stage | Second stage | First stage | Second stage |
| VI Rehabilitation | 0.210*** | | 0.652** | | 0.596** | | 0.594** |
| | (0.0286) | | (0.302) | | (0.253) | | (0.255) |
| DD | | -0.0251*** | | -0.0350*** | | -0.0488*** | |
| | | (0.00917) | | (0.0123) | | (0.0167) | |
| N | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 |
| **Panel C (Home care)** | | | | | | | |
| | | 2SLS | | | | | |
| | | Instrument = linear | | Instrument = square root | | Instrument = log | |
| | OLS | First stage | Second stage | First stage | Second stage | First stage | Second stage |
| VI Home care | 0.0776*** | | 0.379*** | | 0.288*** | | 0.315*** |
| | (0.0132) | | (0.0829) | | (0.0752) | | (0.0782) |
| DD | | -0.0795*** | | -0.0971*** | | -0.136*** | |
| | | (0.0146) | | (0.0148) | | (0.0221) | |
| N | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 |

*Notes*: This table reports the estimation results of equation (1). The OLS column report OLS estimates of VI dummy, $\beta_S$. For the 2SLS estimation, we report three different results using different instruments: linear, square root, and log of DD. The first stage columns report the estimated coefficients of the instruments. The second stage columns report the estimated coefficient of the VI dummy, $\beta_S$. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

Table 5a: Estimation results for "service substitution"

| | Daycare | | Rehabilitation | | Home care | |
|---|---|---|---|---|---|---|
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| VI Daycare | 0.101*** | 0.294** | -0.0468*** | -0.0698 | -0.0768*** | -0.319* |
| | (0.0149) | (0.119) | (0.00962) | (0.105) | (0.0133) | (0.163) |
| VI Rehabilitation | -0.153*** | -0.0941 | 0.189*** | 0.488* | -0.0688*** | -0.475 |
| | (0.0267) | (0.257) | (0.0282) | (0.253) | (0.0175) | (0.304) |
| VI Home care | -0.0265* | -0.134 | -0.000600 | -0.0723 | 0.101*** | 0.412*** |
| | (0.0155) | (0.127) | (0.0115) | (0.0875) | (0.0135) | (0.112) |
| Cragg-Donald F-stat | | 24.19 | | 24.19 | | 24.19 |
| N | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 |

*Notes*: This table reports the estimation results of equation (2). The name of services in each column corresponds to that of the dependent variable (whether the service is included in a care plan), while the service names in rows correspond to the independent variables (whether the CM vertically integrates daycare, rehabilitation, and home care). For brevity, we report the estimated coefficients $\beta_k$ only. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

Table 5b: Estimation results for "service substitution" (By dementia level)

| **Panel A (High Dementia Score)** | | | | | | |
|---|---|---|---|---|---|---|
| | Daycare | | Rehabilitation | | Home care | |
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| VI Daycare | 0.114*** | 0.405** | -0.0535*** | -0.0184 | -0.0843*** | -0.419** |
| | (0.0171) | (0.176) | (0.0113) | (0.129) | (0.0154) | (0.200) |
| VI Rehabilitation | -0.163*** | -0.0470 | 0.202*** | 0.533* | -0.0614*** | -0.272 |
| | (0.0314) | (0.364) | (0.0313) | (0.299) | (0.0191) | (0.361) |
| VI Home care | -0.0363** | -0.148 | 0.000123 | -0.105 | 0.120*** | 0.450*** |
| | (0.0173) | (0.152) | (0.0126) | (0.105) | (0.0161) | (0.139) |
| Cragg-Donald F-stat | | 12.84 | | 12.84 | | 12.84 |
| N | 17,988 | 17,988 | 17,988 | 17,988 | 17,988 | 17,988 |
| **Panel B (Low Dementia Score)** | | | | | | |
| | Daycare | | Rehabilitation | | Home care | |
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| VI Daycare | 0.0702*** | 0.166 | -0.0334*** | -0.133 | -0.0608*** | -0.225 |
| | (0.0162) | (0.139) | (0.0110) | (0.127) | (0.0142) | (0.202) |
| VI Rehabilitation | -0.122*** | 0.0440 | 0.162*** | 0.384 | -0.0850*** | -0.880** |
| | (0.0254) | (0.322) | (0.0284) | (0.290) | (0.0213) | (0.415) |
| VI Home care | -0.00810 | -0.0411 | -0.00401 | -0.0682 | 0.0672*** | 0.351** |
| | (0.0160) | (0.128) | (0.0127) | (0.101) | (0.0150) | (0.148) |
| Cragg-Donald F-stat | | 10.81 | | 10.81 | | 10.81 |
| N | 10,540 | 10,540 | 10,540 | 10,540 | 10,540 | 10,540 |

*Notes*: This table reports the estimation results of equation (2) by dementia level. Panel A (B) reports results for higher (lower) dementia recipients defined by the level of dementia dependency greater than or equal to 2 (lower than 2). The name of services in each column corresponds to that of the dependent variable (whether the service is included in a care plan), while the service names in rows correspond to the independent variables (whether the CM vertically integrates daycare, rehabilitation, and home care). For brevity, we report the estimated coefficients $\beta_k$ only. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

Table 6: Estimation results for "service volume"

| | Daycare | | Rehabilitation | | Home care | | Total of three services | |
|---|---|---|---|---|---|---|---|---|
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| VI Daycare | 837.7*** | 4,946*** | -209.6*** | 265.9 | -740.0*** | -2,603 | -111.8 | 2,609 |
| | (144.6) | (1,760) | (57.08) | (724.4) | (220.7) | (2,024) | (208.1) | (1,774) |
| VI Rehabilitation | -964.3*** | 1,373 | 1,237*** | 3,704** | -545.8** | -3,530 | -272.7 | 1,547 |
| | (196.1) | (3,609) | (181.8) | (1,710) | (256.7) | (3,710) | (266.8) | (3,715) |
| VI Home care | -408.2*** | -4,883** | -26.11 | -639.9 | 907.6*** | 4,300*** | 473.2** | -1,223 |
| | (154.9) | (2,368) | (68.03) | (539.1) | (242.0) | (1,575) | (220.2) | (1,895) |
| Cragg-Donald F-stat | | 24.19 | | 24.19 | | 24.19 | | 24.19 |
| N | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 | 28,528 |

*Notes*: This table reports the estimation results of equation (2), with spending on each service (and its total) as the outcome. The name of services in each column corresponds to that of the dependent variable (how much the service is included in a care plan), while the service names in rows correspond to the independent variables (whether the CM vertically integrates daycare, rehabilitation, and home care). For brevity, we report the estimated coefficients $\beta_k$ only. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

Table 7: Estimation results for "service volume" (By dementia level)

| **Panel A (High Dementia Score)** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Daycare | | Rehabilitation | | Home care | | Total of three services | |
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| VI Daycare | 1,035*** | 7,566** | -267.0*** | 272.5 | -904.9*** | -5,771* | -136.7 | 2,068 |
| | (189.9) | (3,045) | (70.19) | (900.8) | (275.3) | (3,142) | (245.5) | (2,390) |
| VI Rehabilitation | -1,116*** | 4,925 | 1,320*** | 4,158** | -525.8* | -5,194 | -321.5 | 3,889 |
| | (254.0) | (5,942) | (199.1) | (2,116) | (300.4) | (5,853) | (313.5) | (5,083) |
| VI Home care | -523.4*** | -6,324** | -24.66 | -682.1 | 1,260*** | 6,758*** | 711.5*** | -247.5 |
| | (194.7) | (2,947) | (73.33) | (679.2) | (297.5) | (2,337) | (254.5) | (2,115) |
| Cragg-Donald F-stat | | 12.84 | | 12.84 | | 12.84 | | 12.84 |
| N | 17,988 | 17,988 | 17,988 | 17,988 | 17,988 | 17,988 | 17,988 | 17,988 |
| **Panel B (Low Dementia Score)** | | | | | | | | |
| | Daycare | | Rehabilitation | | Home care | | Total of three services | |
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| VI Daycare | 421.1*** | 1,108 | -117.8** | 218.4 | -454.1*** | 1,241 | -150.8 | 2,567* |
| | (92.10) | (979.0) | (59.08) | (821.7) | (128.6) | (1,314) | (139.6) | (1,514) |
| VI Rehabilitation | -555.6*** | -848.6 | 1,092*** | 2,753 | -539.0*** | -1,205 | -2.141 | 700.0 |
| | (128.1) | (2,145) | (178.2) | (1,819) | (170.2) | (2,784) | (195.9) | (3,301) |
| VI Home care | -170.0* | -1,785 | -33.97 | -750.2 | 293.4* | 1,326 | 89.37 | -1,209 |
| | (97.23) | (1,359) | (76.46) | (580.0) | (150.7) | (925.8) | (153.7) | (1,355) |
| Cragg-Donald F-stat | | 10.81 | | 10.81 | | 10.81 | | 10.81 |
| N | 10,540 | 10,540 | 10,540 | 10,540 | 10,540 | 10,540 | 10,540 | 10,540 |

*Notes*: This table reports the estimation results of equation (2) by dementia level, with spending on each service (and its total) as the outcome. Panel A (B) reports results for higher (lower) dementia recipients defined by the level of dementia dependency greater than or equal to 2 (lower than 2). The name of services in each column corresponds to that of the dependent variable (how much the service is included in a care plan), while the service names in rows correspond to the independent variables (whether the CM vertically integrates daycare, rehabilitation, and home care). For brevity, we report the estimated coefficients $\beta_k$ only. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

Table 8: Summary statistics (efficiency effects)

| | Mean | SD | Min | Max | N |
|---|---|---|---|---|---|
| VI any3 | 0.785 | 0.411 | 0 | 1 | 8,062 |
| VI daycare | 0.506 | 0.500 | 0 | 1 | 8,062 |
| VI rehabilitation | 0.142 | 0.349 | 0 | 1 | 8,062 |
| VI home care | 0.527 | 0.499 | 0 | 1 | 8,062 |
| Differential distance | -0.376 | 0.711 | -5.012 | 1.108 | 8,062 |
| Age | 80.931 | 6.919 | 65 | 102 | 8,062 |
| Female | 0.629 | 0.483 | 0 | 1 | 8,062 |
| Coins 0% | 0.069 | 0.254 | 0 | 1 | 8,062 |
| Coins 10% | 0.836 | 0.370 | 0 | 1 | 8,062 |
| Coins 20% | 0.090 | 0.286 | 0 | 1 | 8,062 |
| Coins 30% | 0.009 | 0.092 | 0 | 1 | 8,062 |
| Care level 1 | 0.597 | 0.491 | 0 | 1 | 8,062 |
| Care level 2 | 0.230 | 0.421 | 0 | 1 | 8,062 |
| Care level 3 | 0.097 | 0.296 | 0 | 1 | 8,062 |
| Care level 4 | 0.060 | 0.238 | 0 | 1 | 8,062 |
| Care level 5 | 0.017 | 0.128 | 0 | 1 | 8,062 |
| Care minutes | 52.245 | 20.133 | 32 | 154 | 8,062 |
| ADL | 3.941 | 1.610 | 0 | 7 | 8,062 |
| IADL | 3.107 | 1.121 | 0 | 4 | 8,062 |
| Dementia | 2.541 | 0.889 | 1 | 6 | 8,062 |

*Notes*: This table reports the summary statistics of the health estimation samples. “VI any3” represents a dummy variable for vertical integration, which takes the value of one if recipient’s care manager has any of the following three services under the same ownership, i.e., adult daycare, home care, or rehabilitation service, and zero otherwise. N represents the number of recipients.

Table 9: Balance tests (efficiency effects)

| | DD < median | DD >= median | Difference | p-value |
|---|---|---|---|---|
| Age | 81.011 | 80.860 | 0.151 | 0.328 |
| Female | 0.629 | 0.629 | 0.000 | 0.975 |
| Coins 0% | 0.064 | 0.073 | -0.009 | 0.107 |
| Coins 10% | 0.851 | 0.824 | 0.027 | 0.001 |
| Coins 20% | 0.082 | 0.097 | -0.015 | 0.022 |
| Coins 30% | 0.007 | 0.010 | -0.002 | 0.276 |
| Care level | 1.694 | 1.650 | 0.044 | 0.046 |
| Care minutes | 52.767 | 51.779 | 0.988 | 0.028 |
| ADL | 3.994 | 3.893 | 0.101 | 0.005 |
| IADL | 3.114 | 3.102 | 0.012 | 0.638 |
| Dementia | 2.558 | 2.525 | 0.033 | 0.093 |
| | | | | |
| N | 3,797 | 4,265 | | |

*Notes*: This table reports the user attributes below and above the median of the differential distance for the baseline efficiency sample. We also report the p-value of the t-test for the difference between the means. N represents the number of recipients.

Table 10: Estimation results for efficiency effects (after 14 months)

| | Care minutes | | | ADL | | | IADL | | | Dementia | | | Deterioration of care level | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage |
| VI (Any three services) | | -0.951 | -2.363 | | 0.474 | 0.00974 | | 0.324 | 0.257 | | -0.0573 | -0.0747 | | 0.0408 | 0.0112 |
| | | (5.310) | (6.318) | | (0.426) | (0.504) | | (0.344) | (0.389) | | (0.245) | (0.273) | | (0.127) | (0.146) |
| DD | -0.0703*** | | | -0.0703*** | | | -0.0703*** | | | -0.0705*** | | | -0.0700*** | | |
| | (0.0162) | | | (0.0162) | | | (0.0162) | | | (0.0162) | | | (0.0164) | | |
| Quality | | | -5.589 | | | -1.080** | | | -0.269 | | | -0.0998 | | | -0.110 |
| | | | (4.998) | | | (0.442) | | | (0.322) | | | (0.182) | | | (0.120) |
| Cragg-Donald F-stat | | | 20.17 | | | 20.17 | | | 20.17 | | | 20.07 | | | 18.90 |
| N | 8,407 | 8,407 | 8,062 | 8,407 | 8,407 | 8,062 | 8,407 | 8,407 | 8,062 | 8,407 | 8,407 | 8,057 | 8,265 | 8,265 | 7,928 |

*Notes*: This table reports the estimation results of equation (4) using the baseline efficiency sample. The first row is health outcomes 14 months after the baseline certification. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

Table 11: Estimation results for efficiency effects (after 26 months)

| | Care minutes | | | ADL | | | IADL | | | Dementia | | | Deterioration of care level | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage |
| VI (Any three services) | | 6.040 | 1.940 | | 1.332** | 0.510 | | 0.599 | 0.504 | | 0.193 | 0.0186 | | 0.154 | 0.0166 |
| | | (7.747) | (9.503) | | (0.664) | (0.841) | | (0.502) | (0.632) | | (0.318) | (0.390) | | (0.178) | (0.224) |
| DD | -0.0654*** | | | -0.0654*** | | | -0.0654*** | | | -0.0655*** | | | -0.0638*** | | |
| | (0.0179) | | | (0.0179) | | | (0.0179) | | | (0.0180) | | | (0.0180) | | |
| Quality | | | -6.677 | | | -1.109** | | | -0.168 | | | -0.221 | | | -0.220* |
| | | | (6.209) | | | (0.552) | | | (0.336) | | | (0.209) | | | (0.125) |
| Cragg-Donald F-stat | | | 11.21 | | | 11.21 | | | 11.21 | | | 11.16 | | | 10.63 |
| N | 5,732 | 5,732 | 5,489 | 5,732 | 5,732 | 5,489 | 5,732 | 5,732 | 5,489 | 5,729 | 5,729 | 5,486 | 5,635 | 5,635 | 5,397 |

*Notes*: This table reports the estimation results of equation (4) using the baseline efficiency sample. The first row is health outcomes 26 months after the baseline certification. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

Table 12: Estimation results for efficiency effects (after 14 months)

| | Care munites | | ADL | | IADL | | Dementia | | Deterioration of care level | |
|---|---|---|---|---|---|---|---|---|---|---|
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| VI Daycare | 0.958* | -7.130 | 0.0548 | -0.522 | 0.0377 | -0.0548 | 0.0158 | 0.0265 | 0.0276** | 0.0202 |
| | (0.488) | (6.445) | (0.0387) | (0.473) | (0.0277) | (0.407) | (0.0165) | (0.250) | (0.0110) | (0.167) |
| VI Rehabilitation | -1.110 | -5.357 | -0.0887 | -0.559 | 0.0110 | -0.345 | -0.00663 | 0.0500 | -0.00915 | 0.0563 |
| | (0.736) | (11.05) | (0.0598) | (0.835) | (0.0388) | (0.617) | (0.0252) | (0.428) | (0.0147) | (0.245) |
| VI Home care | -0.307 | -1.340 | -0.0704* | 0.376 | -0.0160 | -0.223 | -0.00150 | -0.0450 | -0.0185 | -0.189 |
| | (0.499) | (6.144) | (0.0417) | (0.470) | (0.0282) | (0.307) | (0.0153) | (0.199) | (0.0116) | (0.142) |
| Cragg-Donald F-stat | | 10.30 | | 10.30 | | 10.30 | | 10.30 | | 9.678 |
| N | 8,407 | 8,407 | 8,407 | 8,407 | 8,407 | 8,407 | 8,407 | 8,407 | 8,265 | 8,265 |

*Notes*: This table reports the estimation results of equation (4) using the baseline efficiency sample. The first row is health outcomes 14 months after the baseline certification. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

Table 13: Estimation results for efficiency effects (after 26 months)

| | Care munites | | ADL | | IADL | | Dementia | | Deterioration of care level | |
|---|---|---|---|---|---|---|---|---|---|---|
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| VI Daycare | 0.752 | -5.006 | 0.0884** | -0.181 | 0.0408 | 0.362 | 0.0327 | 0.381 | 0.0171 | 0.141 |
| | (0.590) | (9.237) | (0.0438) | (0.754) | (0.0336) | (0.567) | (0.0215) | (0.472) | (0.0135) | (0.241) |
| VI Rehabilitation | -0.365 | -14.84 | -0.0634 | -1.477 | 0.000459 | 0.154 | 0.00640 | 0.624 | 0.00145 | 0.372 |
| | (0.938) | (15.87) | (0.0702) | (1.370) | (0.0428) | (0.936) | (0.0312) | (0.798) | (0.0197) | (0.408) |
| VI Home care | -0.207 | 3.927 | -0.0254 | 0.477 | -0.00302 | -0.399 | 0.000429 | 0.0314 | -0.00846 | -0.133 |
| | (0.622) | (6.414) | (0.0462) | (0.521) | (0.0344) | (0.335) | (0.0206) | (0.252) | (0.0134) | (0.123) |
| Cragg-Donald F-stat | | 4.654 | | 4.654 | | 4.654 | | 4.626 | | 4.279 |
| N | 5,735 | 5,732 | 5,735 | 5,732 | 5,735 | 5,732 | 5,732 | 5,729 | 5,638 | 5,635 |

baseline certification. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

Table 14: Estimation results for efficiency effects (after 14 months, by dementia level)

**Panel A (High Dementia Score)**

| | Care minutes | | | ADL | | | IADL | | | Dementia | | | Deterioration of care level | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage |
| VI (Any three services) | | 1.976 | -0.0636 | | 0.824 | 0.430 | | 0.422 | 0.349 | | 0.143 | 0.0681 | | 0.0760 | 0.0300 |
| | | (7.060) | (8.672) | | (0.577) | (0.683) | | (0.348) | (0.400) | | (0.220) | (0.246) | | (0.173) | (0.207) |
| DD | -0.0733*** | | | -0.0733*** | | | -0.0733*** | | | -0.0733*** | | | -0.0723*** | | |
| | (0.0165) | | | (0.0165) | | | (0.0165) | | | (0.0165) | | | (0.0168) | | |
| Quality | | | -11.34* | | | -1.219** | | | -0.323 | | | -0.267 | | | -0.238 |
| | | | (6.172) | | | (0.500) | | | (0.315) | | | (0.190) | | | (0.170) |
| Cragg-Donald F-stat | | | 12.13 | | | 12.13 | | | 12.13 | | | 12.06 | | | 11.14 |
| N | 5,055 | 5,055 | 4,848 | 5,055 | 5,055 | 4,848 | 5,055 | 5,055 | 4,848 | 5,053 | 5,053 | 4,846 | 4,961 | 4,961 | 4,760 |

**Panel B (Low Dementia Score)**

| | Care minutes | | | ADL | | | IADL | | | Dementia | | | Deterioration of care level | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage |
| VI (Any three services) | | -6.493 | -5.295 | | -0.109 | -0.572 | | 0.0763 | -0.000839 | | -0.445 | -0.343 | | 0.00420 | 0.00461 |
| | | (8.375) | (9.423) | | (0.701) | (0.815) | | (0.726) | (0.793) | | (0.429) | (0.460) | | (0.184) | (0.193) |
| DD | -0.0654*** | | | -0.0654*** | | | -0.0654*** | | | -0.0658*** | | | -0.0659*** | | |
| | (0.0218) | | | (0.0218) | | | (0.0218) | | | (0.0218) | | | (0.0220) | | |
| Quality | | | 6.169 | | | -0.681 | | | -0.157 | | | 0.216 | | | 0.123 |
| | | | (7.603) | | | (0.670) | | | (0.610) | | | (0.310) | | | (0.157) |
| Cragg-Donald F-stat | | | 8.107 | | | 8.107 | | | 8.107 | | | 8.086 | | | 7.906 |
| N | 3,352 | 3,352 | 3,214 | 3,352 | 3,352 | 3,214 | 3,352 | 3,352 | 3,214 | 3,349 | 3,349 | 3,211 | 3,304 | 3,304 | 3,168 |

*Notes*: This table reports the estimation results of equation (4) by dementia level using the baseline efficiency sample. Panel A (B) reports results for higher (lower) dementia recipients defined by the level of dementia dependency greater than or equal to 2 (lower than 2). The first row in each panel is health outcomes 14 months after the baseline certification. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

Table 15: Estimation results for efficiency effects (after 26 months, by dementia level)

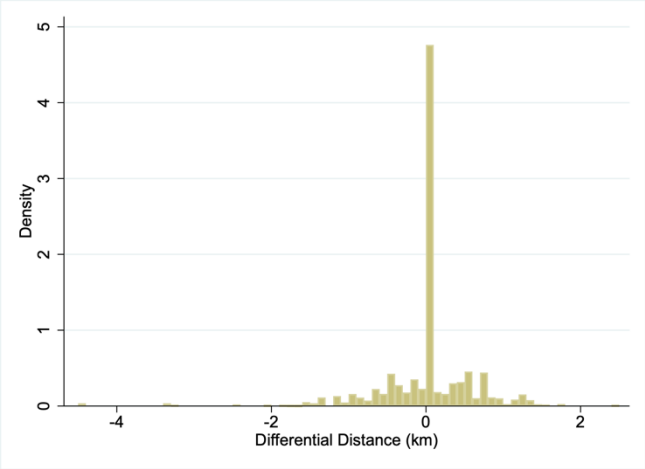
| **Panel A (High Dementia Score)** | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Care minutes | | | ADL | | | IADL | | | Dementia | | | Deterioration of care level | | |
| | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage |
| VI (Any three services) | | 15.49 | 5.582 | | 1.915* | 1.060 | | 0.701 | 0.444 | | 0.445 | 0.207 | | 0.300 | -9.05e-05 |
| | | (11.37) | (14.92) | | (1.001) | (1.189) | | (0.579) | (0.693) | | (0.326) | (0.369) | | (0.271) | (0.375) |
| DD | -0.0651*** | | | -0.0651*** | | | -0.0651*** | | | -0.0650*** | | | -0.0616*** | | |
| | (0.0198) | | | (0.0198) | | | (0.0198) | | | (0.0198) | | | (0.0201) | | |
| Quality | | | -16.62* | | | -1.416** | | | -0.369 | | | -0.358 | | | -0.452** |
| | | | (8.461) | | | (0.650) | | | (0.357) | | | (0.242) | | | (0.192) |
| Cragg-Donald F-stat | | | 6.374 | | | 6.374 | | | 6.374 | | | 6.363 | | | 5.608 |
| N | 3,387 | 3,387 | 3,241 | 3,387 | 3,387 | 3,241 | 3,387 | 3,387 | 3,241 | 3,385 | 3,385 | 3,239 | 3,323 | 3,323 | 3,182 |
| **Panel B (Low Dementia Score)** | | | | | | | | | | | | | | | |
| | Care minutes | | | ADL | | | IADL | | | Dementia | | | Deterioration of care level | | |
| | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage | first stage | second stage | second stage |
| VI (Any three services) | | -5.750 | 3.097 | | 0.811 | 0.0785 | | 0.635 | 0.690 | | -0.205 | 0.211 | | 0.0305 | 0.213 |
| | | (10.48) | (15.19) | | (0.789) | (1.117) | | (0.825) | (1.088) | | (0.529) | (0.679) | | (0.229) | (0.348) |
| DD | -0.0662*** | | | -0.0662*** | | | -0.0662*** | | | -0.0668*** | | | -0.0668*** | | |
| | (0.0236) | | | (0.0236) | | | (0.0236) | | | (0.0236) | | | (0.0236) | | |
| Quality | | | 9.537 | | | -0.660 | | | 0.0317 | | | 0.394 | | | 0.170 |
| | | | (9.305) | | | (0.746) | | | (0.645) | | | (0.379) | | | (0.193) |
| Cragg-Donald F-stat | | | 4.494 | | | 4.494 | | | 4.494 | | | 4.475 | | | 4.747 |
| N | 2,345 | 2,345 | 2,248 | 2,345 | 2,345 | 2,248 | 2,345 | 2,345 | 2,248 | 2,344 | 2,344 | 2,247 | 2,312 | 2,312 | 2,215 |

*Notes*: This table reports the estimation results of equation (4) by dementia level using the baseline efficiency sample. Panel A (B) reports results for higher (lower) dementia recipients defined by the level of dementia dependency greater than or equal to 2 (lower than 2). The first row in each panel is health outcomes 26 months after the baseline certification. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

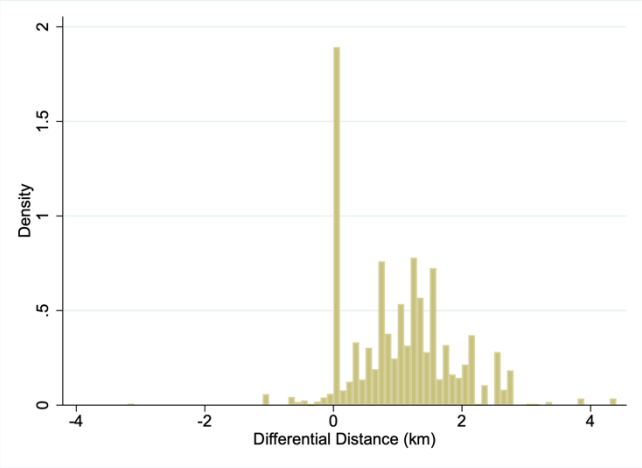
# Figures

Figure 1: Distribution of the DD instrument

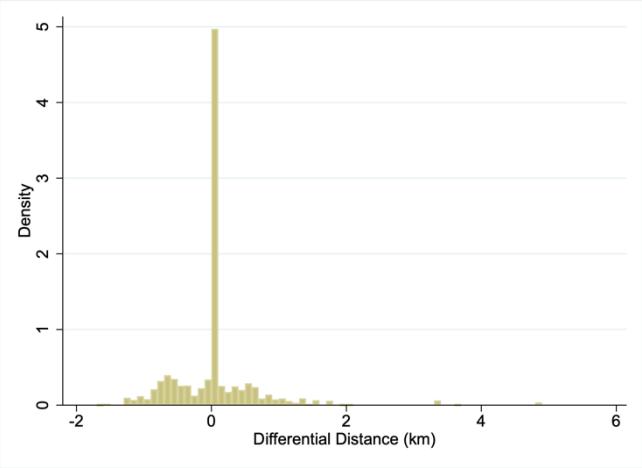
(A) Daycare




(B) Rehabilitation



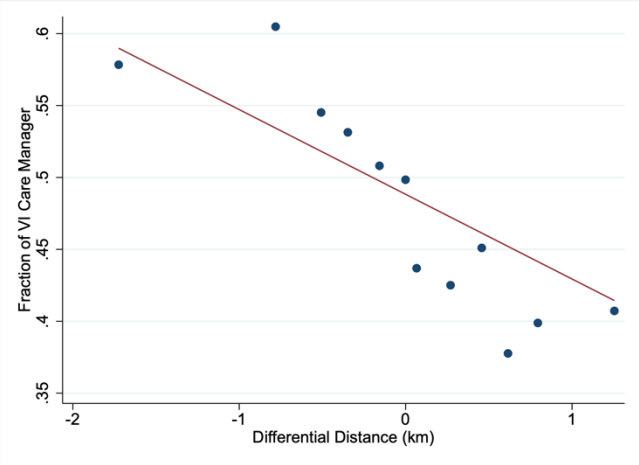

(C) Home care



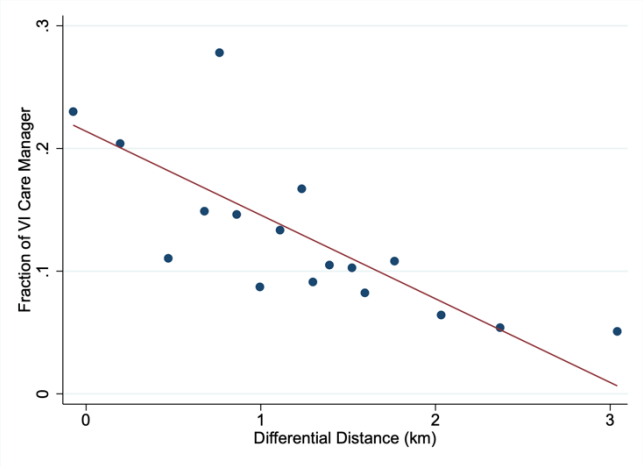

*Notes*: This figure presents the distribution of the differential distance instrument (linear version) for each service. We calculate the distance from the user's residence to the closest care manager who vertically integrates and does not integrate each service and take the difference.

Figure 2: First stage relationship (service choice)

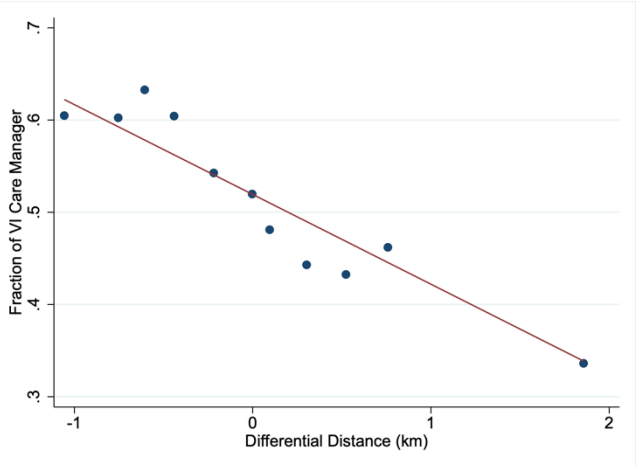
(A) Daycare




(B) Rehabilitation

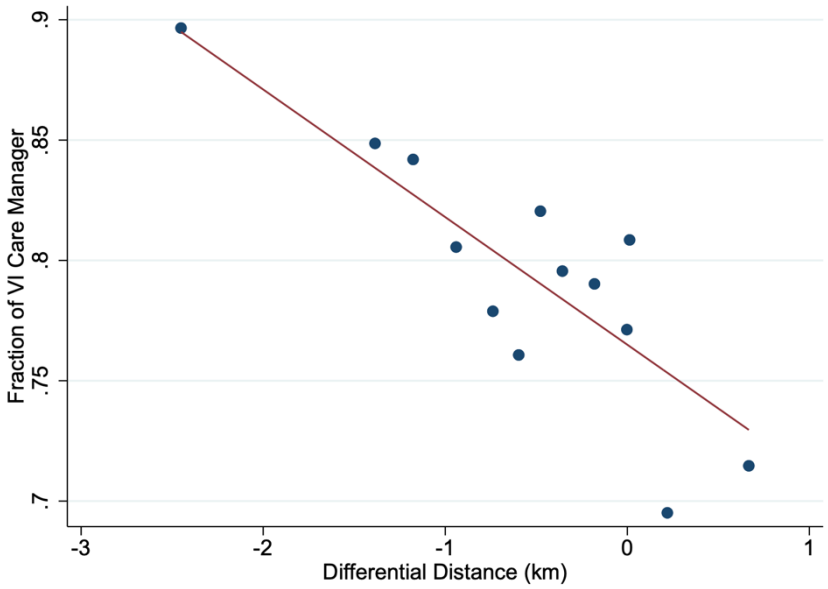



(C) Home care




*Notes*: This figure presents bin scatter plots of the relationship between the DD instrument (linear version) and the vertical integration status of the care managers for each service.

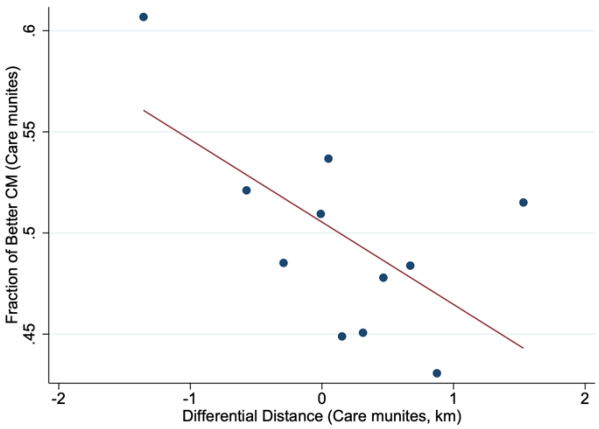
Figure 3: First stage relationship (efficiency, VI CM)



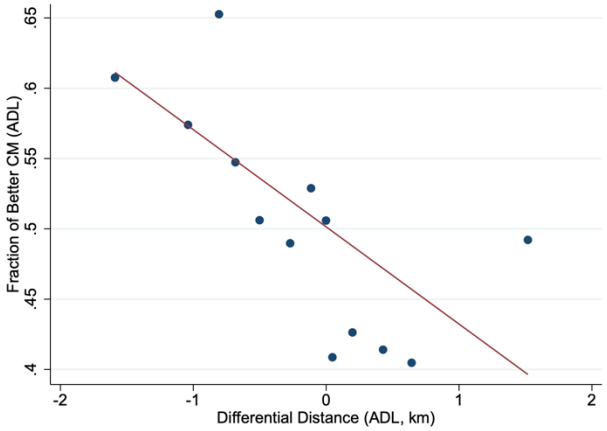

*Notes*: This figure presents bin scatter plots of the relationship between the DD instrument (linear version) and the vertical integration status of the care managers in efficiency estimation sample.

Figure 4: First stage relationship (efficiency, high quality CM)

(A) Care minutes




(B) ADL




*Notes*: This figure presents bin scatter plots of the relationship between the DD instrument for high quality care manager and the quality status of the care managers in efficiency estimation sample.

# Online Appendix

Appendix A: Estimation results for choice/volume with the main sample

Appendix D: Estimation results for sample attrition

# Appendix A (Choice/Volume, Main sample)

Table A1: Estimation results for "service substitution" (By IADL level)

| **Panel A (High IADL Limitation, IADL Score = 4)** | | | | | | |
|---|---|---|---|---|---|---|
| | Daycare | | Rehabilitation | | Home care | |
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| VI Daycare | 0.104*** | 0.371** | -0.0603*** | -0.131 | -0.0712*** | -0.245 |
| | (0.0171) | (0.161) | (0.0110) | (0.137) | (0.0173) | (0.198) |
| VI Rehabilitation | -0.169*** | -0.107 | 0.199*** | 0.545* | -0.0589*** | -0.517 |
| | (0.0273) | (0.338) | (0.0304) | (0.300) | (0.0216) | (0.367) |
| VI Home care | -0.0251 | -0.186 | 0.00557 | -0.0186 | 0.113*** | 0.412*** |
| | (0.0175) | (0.153) | (0.0133) | (0.104) | (0.0173) | (0.138) |
| Cragg-Donald F-stat | | 14.78 | | 14.78 | | 14.78 |
| N | 17,420 | 17,420 | 17,420 | 17,420 | 17,420 | 17,420 |
| **Panel B (Low IADL Limitation, IADL Score < 4)** | | | | | | |
| | Daycare | | Rehabilitation | | Home care | |
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| VI Daycare | 0.0923*** | 0.147 | -0.0273** | 0.00243 | -0.0779*** | -0.363* |
| | (0.0161) | (0.164) | (0.0105) | (0.132) | (0.0124) | (0.190) |
| VI Rehabilitation | -0.133*** | -0.103 | 0.167*** | 0.409 | -0.0703*** | -0.405 |
| | (0.0311) | (0.326) | (0.0291) | (0.298) | (0.0169) | (0.354) |
| VI Home care | -0.0246 | 0.0174 | -0.0107 | -0.179* | 0.0753*** | 0.370*** |
| | (0.0167) | (0.123) | (0.0112) | (0.104) | (0.0121) | (0.133) |
| Cragg-Donald F-stat | | 8.528 | | 8.528 | | 8.528 |
| N | 11,108 | 11,108 | 11,108 | 11,108 | 11,108 | 11,108 |

*Notes*: This table reports the estimation results of equation (2) by IADL level. Panel A (B) reports results for higher (lower) IADL recipients defined by IADL score equal to 4 (lower than 4). The name of services in each column corresponds to that of the dependent variable (whether the service is included in a care plan), while the service names in rows correspond to the independent variables (whether the CM vertically integrates daycare, rehabilitation, and home care). For brevity, we report the estimated coefficients $\beta_k$ only. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

Table A2: Estimation results for "service volume" (By IADL level)

| **Panel A (High IADL Limitation, IADL Score = 4)** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Daycare | | Rehabilitation | | Home care | | Total of three services | |
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| VI Daycare | 982.3*** | 7,403*** | -296.6*** | -516.1 | -873.7*** | -4,081 | -188.0 | 2,806 |
| | (197.5) | (2,541) | (72.85) | (910.3) | (291.6) | (2,683) | (267.1) | (2,230) |
| VI Rehabilitation | -1,164*** | 545.9 | 1,370*** | 3,975* | -586.2* | -4,714 | -381.2 | -193.6 |
| | (247.8) | (5,334) | (219.5) | (2,031) | (321.6) | (5,256) | (340.4) | (4,734) |
| VI Home care | -503.7** | -6,932** | -2.541 | -125.4 | 1,182*** | 6,584*** | 675.9** | -473.2 |
| | (209.5) | (2,962) | (84.29) | (655.7) | (323.3) | (2,114) | (281.9) | (2,267) |
| Cragg-Donald F-stat | | 14.78 | | 14.78 | | 14.78 | | 14.78 |
| N | 17,420 | 17,420 | 17,420 | 17,420 | 17,420 | 17,420 | 17,420 | 17,420 |
| **Panel B (Low IADL Limitation, IADL Score < 4)** | | | | | | | | |
| | Daycare | | Rehabilitation | | Home care | | Total of three services | |
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| VI Daycare | 577.6*** | 1,718 | -89.25* | 1,316 | -497.1*** | -473.6 | -8.737 | 1,718 |
| | (92.36) | (1,909) | (51.55) | (961.5) | (108.1) | (1,365) | (122.6) | (1,909) |
| VI Rehabilitation | -684.2*** | 4,837 | 985.8*** | 3,588* | -374.9*** | -460.0 | -73.37 | 4,837 |
| | (164.8) | (4,016) | (148.3) | (2,038) | (141.4) | (2,538) | (178.5) | (4,016) |
| VI Home care | -193.4** | -559.0 | -59.80 | -1,486** | 433.8*** | 1,140 | 180.6 | -559.0 |
| | (92.34) | (1,203) | (55.33) | (723.8) | (118.7) | (838.6) | (128.3) | (1,203) |
| Cragg-Donald F-stat | | 8.528 | | 8.528 | | 8.528 | | 8.528 |
| N | 11,108 | 11,108 | 11,108 | 11,108 | 11,108 | 11,108 | 11,108 | 11,108 |

*Notes*: This table reports the estimation results of equation (2) by IADL level, with spending on each service (and its total) as the outcome. Panel A (B) reports results for higher (lower) IADL recipients defined by the IADL score equal to 4 (lower than 4). The name of services in each column corresponds to that of the dependent variable (how much the service is included in a care plan), while the service names in rows correspond to the independent variables (whether the CM vertically integrates daycare, rehabilitation, and home care). For brevity, we report the estimated coefficients $\beta_k$ only. N represents the number of recipients. Standard errors are clustered by care manager. ***p<0.01, **p<0.05, *p<0.1.

# Appendix D (Attrition)

Table D1: Sample attrition from the full sample to the main sample

| **Panel (A) Attrition from the full sample to the main sample (Choice/Volume)** | | | | | |
|---|---|---|---|---|---|
| | Attrition dummy | | | | |
| VI Daycare | -0.0687 | | | | -0.0656 |
| | (0.0579) | | | | (0.0621) |
| VI Rehabilitation | | 0.0575 | | | 0.0157 |
| | | (0.0661) | | | (0.0858) |
| VI Home care | | | -0.0332 | | -0.00550 |
| | | | (0.0443) | | (0.0511) |
| VI Any three services | | | | -0.0919 | |
| | | | | (0.0921) | |
| N | 34,591 | 34,591 | 34,591 | 34,591 | 34,591 |
| **Panel (B) Attrition from the full sample to the main sample (Efficiency)** | | | | | |
| | Attrition dummy | | | | |
| VI Daycare | 0.0954 | | | | 0.0832 |
| | (0.0739) | | | | (0.0763) |
| VI Rehabilitation | | -0.0705 | | | -0.0350 |
| | | (0.0789) | | | (0.109) |
| VI Home care | | | 0.0123 | | -0.0249 |
| | | | (0.0484) | | (0.0565) |
| VI Any three services | | | | 0.0233 | |
| | | | | (0.0932) | |
| N | 34,591 | 34,591 | 34,591 | 34,591 | 34,591 |

*Notes*: This table reports how VI status affects sample attrition from the full sample to the main sample. The dependent variable is the indicator, which takes the value one if the observation is dropped and zero otherwise. The right-hand side of the regression is the same as equation (2) or (3) and the same instrumental variable is used. ***p<0.01, **p<0.05, *p<0.1.