Intermediate Care Facilities for Individuals with Intellectual Disabilities: Does Ownership Type Affect Quality of Care?

<u>APPENDIX AND RESULTS PACKET</u>

APPENDIX

Appendix Table of Contents

INTRODUCTION

This Appendix serves as a repository of additional information about our paper, "Intermediate Care Facilities for Individuals with Intellectual Disabilities: Does Ownership Type Affect Quality of Care?" Broadly, it provides additional detail about our data and cleaning procedures; describes how we constructed our alternative datasets and specifications, which are then used to check the robustness of our results; and presents the results of these robustness checks. The document proceeds as follows:

1. In Section 1 (Description of Variable Provenance, Construction, and Cleaning), we describe our Annual Survey, complaints, and deficiency citations data; explain how we identified and constructed the variables that we use in our analysis; and provide in-depth descriptions of our respective cleaning procedures.
2. In Section 2 (Description of Alternative Datasets and Specifications), we describe all of the additional models that we ran in addition to those presented in the paper. Specifically, we outline how and why we created the alternative datasets and specifications that we use to check the robustness of our main results.

3. In Section 3 (Summary Statistics), we present the number (and relative percentage) of residents and facilities by ownership type, as well as summary statistics on the independent and dependent variables in our cleaned dataset, uncleaned dataset, and three alternative datasets, respectively.
4. In the Results Packet following the Appendix, we present the full results (i.e., all coefficients, including on year and state fixed effects) for our preferred specifications, as well as the coefficients for ownership type for all of our robustness check specifications.

Note that although we periodically reference files to help the reader identify where in our replication kit we perform certain procedures,[1] this document does *not* provide a comprehensive overview of the code. Rather, the README_first file—entitled "README_first.docx" in the replication kit—outlines the structure of the replication kit; the inputs, outputs, and function of every R script and do-file used in our analysis; and detailed instructions for how to replicate our work. Note that we also have included a brief introduction to the code, called "Intro to the Code.docx," in the replication kit. This document describes key procedures performed in the code at a relatively high level, and thus serves as an abridged version of (*not* a substitute for) this Appendix and the README_first file.

---

[1] File paths for R scripts and Stata do-files are enclosed in quotation marks. For example, "2_Clean_Ownership_Variable/Code/Clean_Ownershsip_Variable.R" refers to the R script titled "Clean_Ownership_Variable.R," which is located in the "Code" subfolder of the "2_Clean_Ownership_Variable" folder in our replication kit. Similarly, "12_Regressions/2_Analyze_Prepped_Data/Citations.do" refers to the Stata do-file titled "Citations," which is located in the "2_Generate_Tables" subfolder of the "12_Regressions" folder in our replication kit.

1) DESCRIPTION OF VARIABLE PROVENANCE, CONSTRUCTION, AND CLEANING

As is described in the paper, we use three different sources of data, all of which were purchased from the Centers for Medicare & Medicaid Services (CMS) Quality and Inspection Evaluation System (QIES) (CMS, 2017). First, we use CMS Annual Survey data, which are collected as part of the Intermediate Care Facilities for Individuals with Intellectual Disabilities ("ICF-IID," "facility") recertification process and contain all of our model covariates as well as several of our dependent variables (reported nursing ratios, reported direct-care staff ratios, and reported use of drugs, restraint, and time-out rooms to control behavior). Second, we use a file containing data on deficiency citations, which are collected in conjunction with the Annual Survey and can be merged with the Annual Survey data using the "certification date" variable and unique facility-level identifier. Third, we utilize a third file containing data on complaints submitted by third-party stakeholders, which can be merged with the Annual Survey data using the procedure outlined in Section 1.3 (II) below. In all of our regressions, the unit of analysis is the Annual Survey conducted at each facility approximately once per calendar year; we call this unit of analysis the *facility-survey*.

1.1) Defining the Scope of Our Study

I) Annual Survey Data

Our Annual Survey dataset contained information on surveys conducted from May 5, 2006 through April 24, 2017. However, the dataset contained far fewer facility-surveys per year from 2006 through 2008 than from 2009 through 2016. The dataset contained only 1 and 29 facility-surveys for 2006 and 2007, respectively. Moreover, our dataset contained more than 5,000 facility-surveys per year from 2009 through 2016, but only slightly more than 3,000 for 2008.

The puzzling disparity in the number of facility-surveys before and after 2009 led us to suspect strongly that our dataset was incomplete for years before 2009. Comparing the number of observations in our dataset to the number of observations in the public-use Provider of Service data on the CMS website (Center for Medicare & Medicaid Services [CMS], 2021) confirmed these suspicions, while also validating our assumption that the data were complete from 2009 onward. Although the public-use dataset contains only very basic information on each ICF-IID, and only one facility-survey on each ICF-IID per calendar year, it nevertheless contained far more facility-surveys than our dataset for the years 2006-2008. The numbers of observations for later years, however, were very similar. Subsequent correspondence with agency officials did not enable us to pinpoint exactly *why* our data were incomplete prior to 2009 (P. Crawford, Analyst, Internet Quality and Improvement Evaluation System, e-mail correspondence, May 14, 2020; T. Kress, Deputy Director, CMS Division of Quality Systems for Assessment and Surveys, e-mail correspondence, May 21, 2020). However, it did reassure us that the minor discrepancies between our dataset and the public-use versions for 2009 onward were to be expected, since the exact content can vary slightly depending on exactly when the data are pulled and the system from which they are extracted (P. Crawford, Analyst, Internet Quality and Improvement Evaluation System, e-mail correspondence, May 14, 2020). We limited the study to facility-surveys from the years 2009 through 2017, since that is the period for which we are confident that our dataset is complete.

II) Deficiency Citations and Complaints Data

As noted above, because deficiency citations are identified during the ICF-IID recertification process, they can be linked to a specific facility-survey through the "certification date" field and a unique facility-level

identifier. Our deficiency citations dataset includes all deficiency citations recorded during facility-surveys from May 17, 2006, through April 13, 2017. However, since we limited the study to those facility-surveys that took place from 2009 through 2017 (for the reasons explained above), we only consider deficiency citations from those years.

Our complaints dataset contained complaints submitted between October 17, 2004, and April 26, 2017. Unlike deficiency citations, however, complaints can be submitted anytime throughout the year. To estimate our complaint models at the facility-survey level (for which we have data on facility characteristics), we use the technique described in Section 1.3 (II) below to match each facility-survey to complaints that were submitted within a (roughly) six-month window before or after its certification date. Thus, although we only analyze facility-surveys from 2009-2017, we consider complaints from 2008-2017 (inclusive), since a facility-survey occurring in the early months of 2009 could be matched with a complaint submitted in the latter half of 2008.

1.2) Annual Survey Data Drawn from the 3070G Form

The majority of variables used in our analysis are derived from the Intermediate Care Facilities for Individuals with Intellectual Disabilities Survey Report Form CMS-3070G. This form, referred to throughout this Appendix as the "Annual Survey," is collected by state authorities approximately once per calendar year based on information supplied by ICF-IID personnel as part of a facility's yearly recertification (CMS, 2013a, 2018). For the benefit of the interested reader, the Annual Survey is reproduced in its entirety on pages 39-42 of this Appendix (CMS, 2013a).

Yearly recertifications—and in turn, the Annual Surveys—conducted at each facility are not always spaced twelve months apart. The erratic timing of Annual Surveys meant that there was not 1:1 correspondence between calendar years and a given facility's Annual Surveys. Specifically, there were 1223 facilities (about 20% of all facilities in our uncleaned dataset) that had more than one Annual Survey in at least one calendar year. In just over 14% of these 1223 facilities, there is *more than one* calendar year for which more than one Annual Survey is recorded. A similar number of facilities (1221) had no Annual Survey at all in at least one calendar year, and 16% of the latter group had no Annual Survey in at least two calendar years. About ten percent of all facilities in our dataset (606) met both of these criteria. In acknowledgement of this fact, we opted to describe our observations as "facility-surveys" instead of "facility-years" to avoid the misleading implication that Annual Surveys always occur once per calendar year. (We also included calendar year dummies in all of our regressions, which allows us to control for time-variant effects even when there were two facility-surveys in the same calendar year for a given facility.)

At the outset, we spent considerable time investigating the content, structure and reliability of the Annual Survey data. Our analysis brought to light several concerns. First, it was initially unclear which variable in the dataset accurately reflected the number of residents in a given facility, which is included (in some form) as a covariate in all models. Secondly, there were four data quality concerns: the reported ownership type field (our covariate of interest) exhibited telltale signs of reporting and/or coding errors; in a small percentage facility-surveys, a facility reported having more residents than the number of beds on site; adjacent Annual Surveys for a few facilities were separated by suspiciously large gaps; and some facility-surveys reported direct-care staff ratios that were very likely erroneous.

Section I below outlines the process by which we determined which variable in the Annual Survey reflected the facility's number of residents. Section II describes each variable reported in the Annual Survey and its location on the form. Section III describes the various procedures performed to address the four data quality issues listed above. Finally, Section IV describes the limitations of our variable cleaning process and ambiguities in the data that we could not fully address.

I) Identifying the Number of Residents in an ICF-IID

Many of our dependent variables were likely a function of a facility's number of residents. For example, the more residents a facility has, the more care employees, providers, family members, and other stakeholders it is likely to interact with—and thus, *ceteris paribus,* the more complaints one might expect to be submitted against it. Failing to control for a facility's size may therefore inadvertently bias our results.

Our Annual Survey dataset contained three variables that, at first glance, could potentially capture this value: "Client Count: Total – Larger Organization," hereafter referred to as the organization client count; "Client Count: Total – Facility," hereafter referred to as the facility client count; and "Census: Total Residents," hereafter referred to as the resident count. The variable names alone suggest that the resident count (i.e., "Census: Total Residents") represents a facility's number of residents. However, the CMS operations manual for the Annual Survey does not distinguish between "clients" and "residents" (CMS, 2018).[2] For this reason, it was not immediately obvious how the three variables differed and, thus, which one to rely upon in our analysis.

To investigate this question, we examined where each variable was recorded in the Annual Survey, and compared the values for all three variables to each other. This process persuaded us that the resident count was, in fact, the appropriate variable to use for the "number of residents" in our models.

a) *Locating Our Variables on the Annual Survey*

Figures A1 and A2 below depict Sections 16 and 19 of the Annual Survey, respectively, on which we mark the fields corresponding to the three variables of interest. Below these figures, we briefly explain how we identified the correct fields.

Figure A1: Annotated Section 16 (Facility Data) of the Annual Survey

---

[2] With the exception of one provision defining eligibility rules for potential ICF-IID staff members, the operations manual uses "clients" exclusively, with no mention of "residents."

Figure A2: Annotated Section 19 (Individual Characteristics) of the Annual Survey

19. Individual Characteristics
(**NOTE:** The total number in Items B–L (Col.(a)) may exceed the facility's population because some clients have multiple disabilities)

W33: "Total." This is the **resident count** and is equivalent to W36.

Organization client count: If a facility is part of a larger organization, the survey respondent must fill out Section 16, Block B (see Figure A1), which includes field W15, the "Total Number of Clients." In additional Annual Survey instructions given to surveyors, displayed in Figure A3 below, CMS defines W15 as the number of individuals residing in beds owned by the larger organization (CMS, 2013b). The value reported in **W15 is therefore the organization client count**.

Facility client count: The Annual Survey asks for two different client counts (fields W15 and W16 in Figure A1 above). Of these, W16 ("Total Number of ICF/IID Clients") is the only one that *all* facilities—regardless of whether they are part of a larger organization—must complete. The value reported in **W16 is therefore the facility client count**. This inference is further substantiated by the order of the variables in our raw dataset. The Annual Survey contains, in order, W15 (which records the organization client count), W16, and a question about being "community-based." Similarly, in our dataset, the facility's organization client count precedes the facility client count, which in turn precedes a "Community Based Indicator" (W17), once again demonstrating that W16 corresponds to the facility client count.

Resident count: In fields W29-32 (see figure A2 above), respondents must record the respective number of facility residents that fall into four mutually exclusive age categories, and a "Total" in field W33. As the Annual Survey instructions dictate in Figure A4 below, these values refer to *facility* resident characteristics

8

(CMS, 2013b). Because our resident count comes immediately after the four respective age counts in our raw dataset, the "Total" value reported in **W33 is the resident count**. Intuitively (and as mandated by the Annual Survey instructions presented in Figure A4), the numbers entered in fields W33 and W36 must be equivalent (CMS, 2013b). For all facility-surveys in our dataset, the resident count is in fact equal to *both* the sum of the four respective age category counts and the sum of the two sex category counts.

Table A1 (below) summarizes the respective locations of the three variables.

Figure A3: Annotated Annual Survey Instructions Relating to the Organization and Facility Client Counts

Blocks 16 (A-B):	A "Yes" indicates that the CEO directs not only the activities of the ICF/*IID*, but also those of another residential services program (e.g., another ICF/*IID*; another Medicare/Medicaid Provider that serves persons with *ID* regardless of funding source). A "No" indicates that the CEO of the ICF/*IID* does not direct the activity of another residential services program for persons with *IID*. If "Yes" was indicated for 16A, identify the name, address and CEO of the larger organization or agency in 16B (could be the same information for this ICF/*IID* in Block 7. Enter the total bed capacity of all residential services for which the CEO is directly responsible (including the ICF/*IID* bed capacity) in "W14." Do **not** include beds for which the CEO is indirectly responsible. (For example, in some States the CEO of a State-operated institution is also indirectly responsible for **all** beds in a region, including those operated by private providers within that region. Do not include beds directly operated by another agency or organization for the purposes of W14.) Enter the total No. of individuals residing in the beds (including ICF/*IID* individuals) in "W15."
Block 16C:	Enter the No. as requested.

Figure A4: Annotated Annual Survey Instructions Relating to the Resident Count

Blocks 20 (A-L):	INDIVIDUAL CHARACTERISTICS: The last date of the survey is the date by which age is determined. The term "Total" No. refers to the No. of ICF/*IID* individuals fitting the characteristic listed who are currently in the facility.
20 A (1):	Enter the total No. of individuals within each age group regardless of sex.
Blocks 20 A (2):	Enter the No. of individuals by sex and the total. The total should equal the No. entered in 20 (A)(1), Total (W33).

Table A1: Three Resident/Client Variables and Their Respective Locations in the Annual Survey

Value	Variable Name in Raw Data	Section and Block in Annual Survey	Field in Annual Survey
Organization client count	Client Count: Total – Larger Organization	16. Facility Data, Block B (Figure A1)	W15
Facility client count	Client Count: Total – Facility	16. Facility Data, Block C (Figure A1)	W16

Resident count	Census: Total Residents	19. Individual Characteristics, Block A (Figure A2)	W33

b) Identifying Which of the Three Fields Reflects the Number of Facility Residents

Since our analysis is conducted at the level of the individual facility-survey, not the organizational level, we focused attention on two of the three fields in the dataset: the facility client count and the resident count. Since the Annual Survey operations manual does not differentiate between "client" and "resident," one might expect these two values to be equivalent (CMS, 2018). Yet this is not always the case. Upon further inspection, we found three distinct patterns of dubiousness and likely miscoding/misreporting of the facility client count that we did *not* find in the resident count. We explain these patterns in turn:

Pattern 1: In about 11% of facility-surveys, the facility client count is greater than the organization client count. However, neither the Annual Survey nor the accompanying materials specify that the organization count should be restricted to certain types of clients. Moreover, common sense would seem to suggest that the organization client count should include individuals from the *entire* organization, not just a single facility. Thus, it is surprising to observe instances in which the facility count exceeds the organization client count. We inferred that these cases were likely explained by simple miscoding and/or misreporting (or perhaps an inadvertent flipping) of the facility client and/or organization client counts.

Pattern 2: In about 25% of facilities, as is demonstrated in Table A2 below, the facility client count matches the resident count in some years but matches the organization client count in other years. We attribute this puzzling pattern to an ambiguity in the construction of the Annual Survey. Specifically, field W16, which records the facility client count, offers *no explicit guidance* to restrict the reported value to the facility level. As a result, respondents may be inclined to report the number of residents across the larger organization. Furthermore, the facility client count is presented immediately after the organization client count in the survey instrument. The visual proximity of these two fields, and the lack of visually salient guidance on how to interpret the facility client count, introduce significant ambiguity, and likely explain why the facility client count sometimes oscillates between the two other values over time.

Table A2: Facility for which Facility Client Count Switches between Two Other Values

Resident Count	Facility Client Count	Organization Client Count	Certification Date
6	6	191	8/17/2009
6	6	190	8/10/2010
6	6	186	8/17/2011
6	189	189	7/26/2012
6	167	167	7/12/2013
6	6	166	8/8/2014
6	117	117	10/14/2015
6	6	165	10/27/2016

Pattern 3: About 27% of facility-surveys report a facility client count that is *smaller* than the organization client count, yet *larger* than the resident count. The problem is likely that Field W16, which records the facility client count, instructs the respondent to report the "Total Number of ICF/IID Clients," without clarifying that that field should be restricted to the facility level. Yet Field W16 *does* specify that the response should be limited to clients of *ICF-IIDs*. It is possible, then, that survey respondents report the number of ICF-IID residents *across the larger organization* for what we call the facility client count. Since large

organizations may operate facilities other than ICF-IIDs, the number of ICF-IID residents across the larger organization could, in fact, be *smaller* than the organization's total client count.

We demonstrate this pattern in Table A3 below using a facility in our dataset that is part of a larger, five-facility organization. In the first column, we present the facility's resident count reported in each calendar year (i.e., the number of residents in the facility). In the second column, we sum the resident counts of the five component facilities for each year,[3] creating an aggregated resident count for the larger organization. This value roughly represents the number of *ICF-IID* residents across the larger organization (but not necessarily the total number of clients it serves). In every year, the aggregated resident count is identical (or nearly so[4]) to the facility client count. As we suspected, this pattern indicates that respondents were reporting the number of ICF-IID residents served by the larger organization—not just by the ICF-IID being surveyed—as the facility client count, likely due to ambiguity in the Annual Survey.

Table A3: Facility for which the Aggregated Resident Count Roughly Matches the Facility Client Count

Original Resident Count	Aggregated Resident Count	Facility Client Count	Organization Client Count	Calendar Year
8	44	44	83	2008
8	44	44	83	2009
8	43	43	83	2010
8	43	44	83	2011
8	43	44	86	2012
8	44	44	83	2013
8	43	43	82	2014
7	43	43	85	2015
7	43	43	43	2016

In short, very likely due to the ambiguous structure and wording of several portions of the Annual Survey, the facility client count variable is of dubious quality, and is likely susceptible to frequent miscoding. Because the resident count does not exhibit the same patterns of miscoding and/or misreporting, **we included a facility-survey's resident count in our regressions to represent the number residents in a given facility at the time of the Annual Survey.**

---

[3] All facilities in the larger organization have exactly one Annual Survey per calendar year between 2008 and 2016.

[4] The slight deviations between the two values are likely attributable to the fact that facilities are surveyed at different times in a given year. As such, the aggregated resident count may not equal the true aggregated resident count on the date that the individual facility was surveyed.

II) Detailed Description of Variables from the Annual Survey

Table A4: Facility Characteristics Reported in the Annual Survey

Variable	Description
For-Profit Ownership	A binary variable equal to one if the ICF-IID reported that it was owned by a "PRIVATE PROPRIETARY" in field W6 (of the Annual Survey). This is the baseline ownership type, and therefore is the omitted category in our regression models.
Nonprofit Ownership	A binary variable equal to one if the ICF-IID reported that it was owned by a "PRIVATE NON PROFIT" in field W6.
Government Ownership	A binary variable equal to one if the ICF-IID reported that it was owned by either: "CITY/COUNTY" (0.1% of all ICF-IID-years in our dataset); "CITY/TOWN" (0.1%); "COUNTY" (1.4%); or "STATE" (4.0%) in field W6.
Number of Residents	The number of residents reported by the ICF-IID in field W33.
Medium Size Facility Dummy	A binary variable equal to one if the ICF-IID reports between 7-15 residents inclusive in field W33.
Large Size Facility Dummy	A binary variable equal to one if the ICF-IID reports between 16-34 residents inclusive in field W33.
Very Large Size Facility Dummy	A binary variable equal to one if the ICF-IID reports more than 35 residents in field W33.
Direct-Care Staff Per Resident	The number of full-time equivalent direct-care staff reported by the ICF-IID in field W23, divided by the number of residents reported by the ICF-IID in field W33.
Registered Nurses per Resident	The number of full-time equivalent registered nurses reported by the ICF-IID in field W24, divided by the number of residents reported by the ICF-IID in field W33.
Use of Drugs to Control Behavior	A binary variable equal to one if the ICF-IID reports using drugs to control behavior at least once in field W58.
Use of Physical Restraints	A binary variable equal to one if the ICF-IID reports using physical restraint to control behavior at least once in field W59.
Use of Time Out Room	A binary variable equal to one if the ICF-IID reports using a time out room to control behavior at least once in field W60.

Table A5: Demographic Characteristics Reported in the Annual Survey

Variable	Description
Total Residents	The total number of facility residents reported by the ICF-IID in field W33.
Proportion of Residents Under 22	The number of residents under age 22 reported by the ICF-IID in field W29, divided by the number of facility residents reported by the ICF-IID in field W33.
Proportion of Residents Over 65	The number of residents over age 65 reported by the ICF-IID in field W32, divided by the number of facility residents reported by the ICF-IID in field W33.
Proportion Severe/Profound ID	The number of residents with either "severe" or "profound" intellectual disability reported by the ICF-IID in fields W39 and W40 respectively, divided by the number of residents reported by the ICF-IID in field W33.
Proportion Autistic	The number of residents with autism reported by the ICF-IID in field W42, divided by the number of residents reported by the ICF-IID in field W33.
Proportion Cerebral Palsy	The number of residents with cerebral palsy reported by the ICF-IID in field W43, divided by the number of residents reported by the ICF-IID in field W33.
Proportion Epilepsy	The number of residents with either "controlled" or "uncontrolled" epilepsy reported by the ICF-IID in field W46, divided by the number of residents reported by the ICF-IID in field W33.
Proportion Language Impairment	The number of residents with a speech or language impairment reported by the ICF-IID in field W50, divided by the number of residents reported by the ICF-IID in field W33.
Proportion Hearing Impairment	The number of residents with a hearing impairment reported by the ICF-IID in field W53, divided by the number of residents reported by the ICF-IID in field W33.
Proportion Visual Impairments	The number of residents with a visual impairment reported by the ICF-IID in field W56, divided by the number of residents reported by the ICF-IID in field W33.
Proportion Non-ambulatory	The number of residents with a non-ambulatory disability reported by the ICF-IID in field W49, divided by the number of residents reported by the ICF-IID in field W33.
Proportion Medical Care Plan	The number of residents with a medical care plan (requiring licensed nursing care on a 24 hour basis) reported by the ICF-IID in field W57, divided by the number of residents reported by the ICF-IID in field W33.

III) Description of the Variable Cleaning Process

In preparing the datasets for analysis, four concerns came to light. First, the reported ownership type field (our covariate of interest) exhibited telltale signs of reporting and/or coding error. Secondly, in a small percentage of facility-surveys, a facility reported having more residents than beds. Third, adjacent Annual Surveys for a few facilities were separated by suspiciously large gaps, which made it difficult to reliably assign complaints to facility-surveys. Finally, some facility-surveys contained direct-care staff ratios that were very likely erroneous, casting doubt on the validity of reported direct-care staffing ratios. To address these concerns, we implemented four separate cleaning procedures, as explained below:

a) Implausible Ownership Type Patterns

As discussed in the paper, some facility-surveys in our dataset exhibited highly implausible patterns in reported ownership type. For example, some facilities reported a consistent ownership type for years, changed to a different ownership type for a single year, and then reverted back to the original (reported) ownership type for all remaining years in the dataset. In situations like this, we felt it was highly likely that ownership status was misreported and/or miscoded in the single aberrant year.

To address likely miscoding and/or misreporting of our ownership type field, we implemented the following data cleaning procedure:[5]

1) **First, we retained *all* facility-years from facilities that displayed credible patterns of ownership type.** We retained all of the facility-surveys in these facilities in their original form, amounting to 38,729 facility-surveys, or 82% of facility-surveys in our original dataset. We considered two ownership patterns to be credible:
    a. Facilities with more than one facility-survey that recorded the same ownership type for all facility-surveys. (Note that if a facility reported an "other" ownership type for all facility-surveys, it was subsequently dropped from the dataset in step 3.)
    b. Facilities that only reported a single change in ownership type, which occurred *between* the first and last facility-surveys in the dataset. For example, a facility that appears in the dataset from 2008 through 2017, and switched ownership type *once* between 2009 and 2016 (e.g. switched from nonprofit to for-profit status in 2014), and was coded consistently in every year prior and subsequent to this change (e.g. was coded as nonprofit from 2008 through 2013, and coded as for-profit from 2014 through 2017), was deemed facially credible. Importantly, facilities that reported a single change in ownership type, but for whom this change occurred in the very first or very last facility-year in the dataset, was not deemed "facially credible" in this sense. For example, a facility that was coded as nonprofit in 2009 (the first year it entered the dataset), but was coded as for-profit in every subsequent year, did not meet this criterion. (Note again that if a facility reported an ownership type of "other" for more than one facility-survey, it was subsequently dropped from the dataset in step 3, even if it was initially retained in this step.)

2) **Secondly, we identified facilities with a single aberrant facility-survey in which ownership type had likely been miscoded.** Our general approach was to correct or drop the aberrant facility-year, while retaining all remaining facility-surveys in their original form. Throughout this step, we dropped 374 facility-surveys, corrected 436 facility-surveys, and retained 4,660 facility-surveys in their original form, which amounted 0.8%, 0.9%, and 9.9% of all facility-surveys in our original dataset, respectively. More specifically:
    a. For facilities with a single aberrant value of ownership type in an otherwise consistently coded series of facility-surveys, we changed the aberrant value to match the other values in the facility. For example, if a facility appears in the dataset from 2007 through 2013, and is coded as a nonprofit in every single year except 2010 (when it was coded as a for-profit facility), we assumed the value for 2010 had been miscoded, and changed its value to nonprofit.
    b. For facilities whose reported ownership type changed only once, in either the very first or last recorded facility-survey, we dropped the aberrant facility-survey and retained all other facility-surveys. This criterion would be met, for example, by a facility that appears in the dataset from 2007 through 2013, and is coded as a nonprofit in every single year except 2013. In such situations, we removed the aberrant facility-survey from our dataset but retained all other

---
[5] We perform the cleaning procedure for the ownership type field in
"2_Clean_Ownership_Variable/Code/Clean_Ownership_Variable.R."

facility-surveys, since we had no way to know whether the aberrant value represented a true change in ownership type or just a coding error. (So referring again to the example above, we would remove the facility-survey for 2013, but retain all other facility-surveys.) Note that in effect, this exclusion criterion eliminated all facility-surveys from facilities with only two facility-surveys with differing ownership types (e.g. nonprofit, then for-profit), because it flagged the first and last facility-surveys, in turn, as deviating from an otherwise constant sequence.[6]

    c. <u>We dropped facilities with only one facility-survey from the dataset.</u> With no neighboring facility-surveys to use as a basis for comparison, we were unable to assess the credibility of these "singleton" facility-surveys, and therefore opted to remove them to avoid potential miscoding and/or misreporting.

3) **Thirdly, we dropped all facility-surveys that were *retained* in the first two steps from facilities that, after the first two steps, still contained one or more facility-surveys with an "other" ownership type**. Our concern was that facilities coded as "other" for multiple surveys were not comparable to other facilities, since the "other" category is not clearly defined by CMS.[7] Therefore, we decided to drop all facility-surveys for any facilities that, after the first two steps (above) were completed, still contained at least one facility-survey with an "other" ownership type. For example, if a facility was categorized as "other" for five consecutive facility-surveys—or alternatively, was coded as nonprofit for three consecutive surveys and then as "other" for the two remaining surveys—all of its facility-surveys would be retained in the first step of the cleaning procedure, but then removed from the cleaned dataset in Step 3. To take another example, if a facility was categorized as "other" in four consecutive facility-surveys and nonprofit in a fifth (final) facility-survey, the fifth facility-survey would be removed in Step 2, and the four remaining facility-surveys (coded as "other") would be removed in Step 3. Using this procedure, we removed an additional 2,487 facility-surveys from the dataset (5.3% of all facility-surveys in our original dataset).

4) **Lastly, we dropped all facility-surveys for facilities that had not already been dropped or retained in Steps 1-3.** Once the first three steps were completed, the only facilities that remained had not been addressed were ones that displayed a highly erratic or haphazard pattern, fluctuating between different ownership types multiple times within the study period. Although we felt these data were very likely to contain coding errors, there was no principled way for us to pinpoint such errors or devise a method to correct them. Therefore, we dropped these erratic facilities—which comprised 3004 facility-surveys, or 6.4% of all facility-surveys—from the dataset.

Note that because ownership type is the key independent variable in all of our regressions, we performed the above procedure on the "cleaned" dataset used in *all* of our regressions. The uncleaned dataset that we use for robustness checks (see Section 2.1 below), on the other hand, retains the original values of the ownership type field.

---

[6] The removal of facility-surveys from facilities with only two facility-surveys with different ownership types is not this straightforward in the code, where facilities with this ownership pattern are only flagged as having an aberrant *first* facility-survey. That is, the functions we create in "2_Clean_Ownership_Variable/Code/Clean_Ownership_Variable.R" for step 2b only remove the first facility-survey in these facilities (e.g. the nonprofit facility-survey in a two-survey facility that reports being nonprofit in its first year and for-profit in its second year). To ensure that *both* the first and last facility-surveys in these facilities are removed, we also include a function later in "2_Clean_Ownership_Variable/Code/Clean_Ownership_Variable.R" that removes *all* facility-surveys from facilities that exhibited this pattern. The net effect is therefore the removal of *both* facility-surveys in all facilities that exhibited this ownership pattern.

[7] We tried extensively to try to find out what kinds of facilities tended to make themselves as having an ownership type of "other." We were told anecdotally from officials at the California Department of Health that most of the facilities falling under the "other" ownership category are sole proprietorships (C. Egel, California Department of Health, e-mail correspondence, March 14, 2018).

b)   Total Number of Facility Residents Larger than Facility Bed Count

   The Annual Survey contains separate questions for the number of residents and the total number of beds at each facility. Intuitively, a facility's reported number of residents should not exceed the reported number of beds. Nevertheless, our dataset contained 401 facility-surveys (0.8% of all facility-surveys) whose reported number of residents was *greater than* the reported number of beds. Since we could not determine whether this pattern was due to miscoding of the number of residents and/or the number of beds, we dropped these anomalous facility-surveys from the cleaned dataset to help ensure that the residents (used as a covariate in some form in all models) did not contain erroneous values.[8]

c)   Implausibly Large Gaps between Adjacent Facility-Surveys and Adjustments to Complaints Models

   There were 12 facilities in our dataset for which two adjacent Annual Surveys were separated by at least 24 calendar months, raising the concern that they temporarily closed down or experienced other significant disruptions that made them different from other facility-surveys in our dataset, and thus potentially could bias the number of complaints filed against them.
   To address this concern, we flagged the facility-surveys in question[9] and later removed them from the complaints models (in which the dependent variables are total and substantiated complaints, respectively).[10] Using this procedure, we dropped 24 facility-surveys (0.05% of all facility-surveys in our original dataset) from the complaints models.[11] As is explained in Section 2.1 below, we retained these 24 facility-surveys in the "uncleaned" dataset and in all other models (none of which utilize information on complaints). As a practical matter, however, these 24 facility-surveys had virtually no impact on our results since they only affected the complaints models and, even in those models, constituted such a small fraction of the total dataset.

d)   Implausible Reported Direct-Care Staff Ratios and Adjustments to Direct-Care Staff Ratio Models

   We found evidence that the number of reported direct-care staff, which we use to construct the direct-care staff ratio (one of our dependent variables), was frequently miscoded and/or misreported. The number of direct-care staff at a facility is subject to strict statutory requirements based on the number of residents. Staff ratios that fell far above (or far below) these levels seemed highly dubious. For example, if a facility reportedly employed *one-tenth* of the statutorily-required minimum number of staff for a given year, we felt that this value was very likely miscoded.
   To address the likely miscoding and/or misreporting in the reported direct-care staff ratio, we implemented a four-step cleaning procedure that flagged all facility-surveys whose staff ratios seemed extremely improbable, so that they could be omitted from regression models in which the number of direct-care staff per resident was the dependent variable[12]:

1) **Step 1:** 42 C.F.R. § 483.430 (2019) establishes that when clients are present at an ICF, there must be a direct-care staff person on duty on a "24-hour basis." Every facility-survey in our dataset reports at least one resident and is therefore subject to this requirement. In order to translate this statute into a threshold

---

[8] We drop these facility-surveys from our cleaned dataset in "2_Clean_Ownership_Variable/Code/Clean_Ownership_Variable.R."

[9] We flag these 24 facility-surveys in "1_Merge/Code/Merge.R."

[10] We remove the facility-surveys from the complaints models in "12_Generate_Tables/1_Prep_Data/Code/1_Prep_Data.do." We then estimate the models in "12_Generate_Tables/2_Generate_Tables/Code/Generate_Tables_new.do."

[11] "Complaint models" is used throughout the Appendix to refer to models that analyze total complaints, total complaints per resident, substantiated complaints, and substantiated complaints per resident.

[12] We flag the facility-surveys with implausible direct-care staff ratios in "2_Clean_Ownership_Variable/Code/Clean_Ownership_Variable.R" and remove them from our regressions in "12_Regressions/1_Prep_Data/Code/Prep_Data.do."

that would exclude only miscoded facility-surveys—as opposed to marginally non-compliant facilities, or those with exigent circumstances that would allow temporarily for a lower number of FTE direct-care staff—we assumed that in any given facility, *all* of the residents could be offsite at certain times. Specifically, we assumed that all residents could participate in day programs eight hours a day, five days a week, and also spend the entirety of every weekend at home. Under this conservative set of assumptions, there should be a resident (and therefore at least one direct-care staff member) onsite for 80 hours of the week, amounting to a floor of 2 FTE direct-care staff. **Using this logic, we excluded all facility-surveys whose number of FTE direct-care staff fell at or below 2.0 from our direct-care staff ratio regression.**

2) **Step 2:** 42 C.F.R. § 483.430 (2019) also establishes thresholds for direct-care staff *ratios* based on the level of care necessary for the facility's resident population. To calculate a *floor* for credible staff ratios, we assume that all residents in the facility are "within the range of mild retardation" and therefore require the statute's *least stringent ratio* of only 1 direct-care staff member for every 6.4 residents. (We further assume, as explained above, that all residents are only onsite 80 hours a week.) Taken together, these assumptions imply that each facility-survey should report a direct-care staff ratio of at least 0.3125 to comply with the statute.[13] **Using this logic, we dropped all facility-surveys whose staff ratio fell at our below 0.3125 from our direct-care staff ratio regression.**

3) **Step 3:** To establish a *ceiling* for credible direct-care staff ratios, we also relied on 42 C.F.R. § 483.430 (2019). To translate the statute into a ceiling for credible direct care-staff ratios, we assume that all residents in the facility may be present at *all* times and are "children under the age of 12, severely and profoundly retarded clients, clients with severe physical disabilities, or clients who are aggressive, assaultive, or security risks, or who manifest severely hyperactive or psychotic-like behavior," and therefore require the *highest* direct-care staff-to-resident ratio, which equates to a requirement of 1.3125 FTE direct-care staff per resident.[14] To ensure that we do not exclude facilities that voluntarily choose to exceed the statutory requirements, **we exclude facility-surveys with direct-care staff ratios that were *more than three times as high* as the most stringent staffing ratio requirement (i.e., a ratio of 3.94 or above).**

4) **Step 4:** Finally, in what we call the "intra-facility cleaning step," we dropped facility-surveys in which the reported direct-care staffing ratio seems improbably high or low relative to its neighbors *and* the median reported direct-care staffing ratio for that facility.[15] We carry out this procedure in two steps:
   a) **First, we flag all facility-surveys whose reported direct-care staffing ratio is strictly greater than twice (or strictly less than half) the reported direct-care staffing ratios in both neighboring facility-surveys *and* the median reported direct-care staffing ratio**

---

[13] For the purposes of calculating this staff ratio, we consider a hypothetical facility with one resident. As required by statute, a direct-care staff member would need to be at the facility 1 staff hour for every 6.4 hours the resident is on sight. Under the assumptions we use to construct the floor, a resident would *only* be on site for 80 hours a week. The statute would therefore require at least 12.5 staff hours per week (as 12.5 ÷ 80 is equivalent to 1 ÷ 6.4). Because the number of FTE direct-care staff is defined as the number of staff working 40 hours a week, we divide 12.5 hours by 40 hours to achieve 0.3125, the staff ratio necessary to comply with the statute's *least stringent* requirement.

[14] For the purposes of calculating the ratio, we consider a hypothetical facility with one resident. As required by statute, a direct-care staff member would need to be at the facility 1 staff hour for every 3.2 hours the resident is on sight. Under the assumptions we use to construct a ceiling, a resident is on sight at *all times*, or 168 hours a week. This means that the statute would require at least 52.5 staff hours per week (as 52.5 ÷ 168 is equivalent to 1 ÷ 3.2). Because the number of FTE direct-care staff is defined as the number of staff working 40 hours a week, we divide 52.5 hours by 40 hours to achieve 1.3125, the staff ratio necessary to comply with the statute's *most stringent* requirement.

[15] Note that we performed the cleaning procedure for the direct-care staff ratio *after* we removed facility-years with an implausible reported ownership type or reported number of residents from our dataset. We therefore do *not* consider facility-surveys removed in these two cleaning procedures when checking if a facility-survey's staff ratio is aberrant relative to others in the facility.

among all facility-surveys for that facility. If we were to discard all facility-surveys meeting this criterion, however, we might inadvertently discard some facility-surveys in which there was a temporary (one-year) spike or drop in the staffing ratio due to a corresponding temporary drop (or spike) in the number of residents, and in which the number of direct-care staff could not be adjusted instantaneously in response to this change. In other words, we do not want to discard facility-surveys whose staffing ratios merely *seem* implausible because of short-term "stickiness" in the number of personnel.

b) To guard against the latter possibility, in the second step, we only discard facility-surveys with improbably high or low staffing ratios (i.e., that met the first criterion) in which the aberrant observation could *not* reasonably be explained by a contemporaneous spike or drop in the number of residents. Specifically, **if we observed an improbably large one-year spike in the direct-care staffing ratio for a given facility-survey, *we only* discarded that facility survey if the number of residents for the corresponding facility-survey was no more than 40% lower than the neighboring and median number of reported residents for that facility. Conversely, if we observed an improbably large one-year drop in the direct-care staffing ratio for a given facility-survey, we only discarded that facility survey if the number of residents for the corresponding facility-survey exceeded that of the median and neighboring facility-surveys by no more than 80%.** Using this two-step procedure, we strike a balance between dropping observations with highly implausible staffing ratios, while accounting for the possibility that staffing ratios may exhibit short-term "stickiness" because of the difficulty of immediate hiring (or laying off) staff in response to a large deviation in the number of residents.

Importantly, we implemented Steps 1-3 before Step 4 to ensure that erroneous data did not prevent a facility-survey from being flagged for having an aberrant spike or drop in its direct-care staff ratio relative to adjacent facility-surveys and the median direct-care staff ratio across the facility. Table A6, below, demonstrates the importance of completing Steps 1-3 before Step 4 using a real facility from our dataset. Without first removing the erroneous facility-surveys that reported less than 2 FTE direct-care staff (the first four facility-surveys) in Step 1, the facility's median direct-care staff ratio would be biased *downward* to 0.4, and we would thus fail to flag the facility's 2012 direct-care staff ratio as aberrant relative to the facility's median.

Table A6: Facility Demonstrating the Need to Implement Steps 1-3 before Step 4

Certification Date	Reported Number of FTE Direct-Care Staff	Reported Number of Total Residents	Direct-Care Staff Ratio
9/26/2008	0.1	5	0.02
9/18/2009	0.1	5	0.02
1/10/2011	0.1	5	0.02
1/19/2012	0.1	6	0.017
10/25/2012	2	5	0.4
10/25/2013	10	5	2
12/15/2014	11	5	2.2
10/6/2015	8	5	1.6
12/14/2016	10	5	2

It is also important to note that we only flagged facility-surveys whose staff ratios were aberrant relative to *both* the facility's median *and* the ratios reported in the adjacent facility-surveys, because applying either criterion alone could have generated false positives. Table A7 below illustrates this point using an example facility from our dataset. The facility's median staff ratio is 1.65, so checking if the staff ratios were aberrant

relative to the median alone (without also checking if they were aberrant relative to the adjacent direct-care staff ratios) would have improperly eliminated the first five facility-surveys, despite there being no indication that they were miscoded or misreported.

Table A7: Facility Demonstrating the Need to Test for Aberrance Relative to Adjacent Facility-Surveys and Median Direct-Care Staff Ratio across Entire Facility

Certification Date	Reported Number of FTE Direct-Care Staff	Reported Number of Total Residents	Direct-Care Staff Ratio
8/21/2008	3.3	7	0.47
8/26/2009	3.3	7	0.47
8/18/2010	3.3	7	0.47
09/09/2011	3.3	6	0.55
08/14/2012	11.6	7	1.66
08/07/2013	11.75	7	1.68
08/21/2014	11.6	7	1.66
08/05/2015	11.5	7	1.64
08/17/2016	11.6	7	1.66

Finally, Table A8 below illustrates the importance of accounting for potential "stickiness" in staffing levels. In response to a sudden surge in residents, a facility may have difficulty hiring more direct-care staff, leading to a temporary drop in the direct-care staff ratio. Conversely, if a facility suddenly loses a number of residents, it may opt to keep direct-care staff on its payroll in the hopes that the facility may soon fill empty beds, leading to a temporary *spike* in the direct-care staff ratio. It is clear from the table that the facility's direct-care staff ratio appears aberrant in 2012 because it experienced a dramatic drop in number of residents, yet opted to retain almost half of its staff. As explained above, we do *not* remove the 2012 facility-survey from the dataset, because the large, precipitous spike in the direct-care staff ratio in 2012 coincided with a similarly large, precipitous fall in the number of residents.

Table A8: Facility for which a Temporary Direct-Care Staff Ratio Spike is Driven by a Drop in Residents

Certification Date	Reported Number of FTE Direct-Care Staff	Reported Number of Total Residents	Direct-Care Staff Ratio
09/23/2008	8.25	5	1.65
08/11/2009	8.25	5	1.65
09/14/2010	8.25	5	1.65
09/27/2011	8.25	5	1.65
09/26/2012	3.75	1	3.75
09/16/2013	5.50	3	1.83
08/14/2014	5.60	4	1.4
09/10/2015	8.25	5	1.65
09/02/2016	8.25	5	1.65

As demonstrated by Table A9, below, the vast majority of facility-surveys that were dropped from the cleaned dataset used in the direct-care staff models were dropped in Steps 1-3 of the four-step procedure:

Table A9: Share of Facility-Surveys Dropped from the Direct-Care Staff Ratio Regressions at Each Stage of the Cleaning Process

	Full Dataset	Gov.	Non-Profit	For-Profit
Results of Statute Based Cleaning (Step)				
% obs. with < 2 FTE direct-care staff (1)	1.67%	1.12%	0.95%	2.70%
% obs. with direct-care staff ratio < .3125 (2)	1.82%	1.36%	1.07%	2.88%
% obs. with direct-care staff ratio > 3.9375 (3)	1.11%	3.27%	1.03%	0.92%
Total % facility-surveys dropped in Steps 1-3*	2.99%	4.63%	2.12%	3.91%
Results of Intra-Facility Cleaning Step				
% facility-surveys dropped in Step 4	0.35%	0.33%	0.21%	0.54%
Results of Complete Staff Cleaning Procedure				
Total % facility-surveys dropped	3.34%	4.96%	2.33%	4.44%

* Note that the values in the total values in this row are not equal to the sums of the percentages listed in the first three rows, because there is some redundancy in the first three steps (i.e., different steps of the cleaning procedure sometimes remove the *same* facility-surveys).

e) Order of Implementation for Four Cleaning Procedures

To carry out the various procedures outlined in this section (cleaning facility-surveys with miscoded ownership types and resident counts; cleaning widely-spaced facility-surveys for purposes of the complaints models; and cleaning direct-care staff ratios for purposes of the direct-care staff ratio models), we used the following sequence:

First, we flagged widely spaced facility-surveys. We performed this procedure at the outset to ensure that we only flagged facility-surveys whose gaps with adjacent facility-surveys were *not* caused by the removal of facility-surveys in our other cleaning procedures.

Secondly, we removed (or modified) the facility-surveys with implausible ownership types.

Third, we removed all facility-surveys from our dataset whose reported resident count was larger than the reported bed count. (Note that if we removed facility-surveys with dubious resident counts *before* those with dubious ownership types, we would have been unable to analyze *all* reported ownership type values across time for each facility.)

Finally, we performed the four-step procedure for cleaning the direct-care staff variable.

As noted above and discussed further in Section 2.1, facility-surveys with implausible ownership types and/or resident counts were removed from the cleaned dataset used for *all* models. In contrast, facility-surveys with suspiciously large gaps (casting doubt on the accuracy of the complaints variable) or implausible direct-care staff ratios were *only* removed from the cleaned datasets in which complaints and direct-care staff ratios were used (respectively) as dependent variables.

IV) Limitations of the Variable Cleaning and Double Counting Concerns

As explained in Section 1.1 (II) above, we calculated a facility-survey's full-time equivalent (FTE) registered nursing staff ratio ("nursing ratio") by dividing the number of FTE registered nurses by the number of facility residents, both of which are reported in the Annual Survey. Although we attempted to be as thorough as possible in our variable cleaning process, we could not clean the nursing ratio in the same manner as the direct-care staff ratio because the relevant regulations (42 C.F.R. § 483.450, 2019) do not set any minimum ratios for nurses, but rather specify that they must comply with an individual resident's medical care plan. Therefore, nursing ratios can, in theory, change dramatically from year to year in response to resident turnover

and/or changes in medical care plans. Given these considerations, we do not remove any facility-surveys from our regressions based on patterns observed in the reported number of registered nurses.

Importantly, textual ambiguity in the Annual Survey and accompanying instructions introduce some possibility of "double counting" in the reporting of direct-care staff and registered nurses (CMS, 2013b). The guidance for entering the number of FTE direct-care staff (field W17A) instructs the surveyor to include "staff who provide direct-care services to individuals at their living units" (CMS, 2013b). This language could include registered nurses serving in a direct-care capacity. Therefore, in theory, a single nurse could be counted in *both* the FTE direct-care staff *and* FTE registered nurses categories. As illustrated by Table A10, our data suggests that double counting is likely present in a small number of facility-surveys. Specifically, in 1.28% of facility-surveys, the sum of FTE registered nurses, FTE licensed nurses, and FTE direct-care staff is *greater* than the number of total FTE staff — a value that is supposed to include all three of the latter categories in addition to support staff not involved in resident care (CMS, 2013a). Notably, concerns about the possibility of double counting have been raised in prior literature (Karon & Beutel, 2001).

Table A10: Breakdown of Double Counting in Staff Variables

	Full Dataset	Government	Non-Profit	For-Profit
# of Facility-Surveys with Double Counting	510	8	163	339
% of Facility-Surveys with Double Counting	1.28%	0.34%	0.75%	2.13%

1.3) Complaints and Deficiency Citations Data

I) Provenance of Complaints and Deficiency Citations

The remaining two datasets—a deficiency citations dataset and a complaints dataset—are compiled by state authorities from information obtained from independent third parties. Regulatory deficiencies are assessed during unannounced inspections conducted by state agency officials or CMS regional office representatives in conjunction with the Annual Survey, and both the inspections and surveys are part of the facility's yearly recertification process (CMS, 2018). Complaints against ICF-IIDs, on the other hand, are submitted by concerned stakeholders—such as consumers, family members, advocates, or health care providers—to state agencies or CMS regional offices throughout the year. Both deficiencies and complaints are recorded in CMS'S ASPEN system (CMS, 2018, 2019).

II) Deficiency Citations and Complaints Variable Construction

As is explained in the "Methods" section of the paper, we used both the frequency and per-resident rates of deficiency citations, total complaints, and substantiated complaints, respectively, as dependent variables in our regressions. In these models, we sought to control for the characteristics of the facility—including ownership type, our covariate of interest; number of residents, which is used as a covariate and for constructing the per-resident rate dependent variables; and resident characteristics—at the time the complaint or deficiency citation was filed. The difficulty was that complaints and deficiency citations datasets contained no facility-level information whatsoever besides the date of the complaint/citation and a unique facility-level identifier. The only way to estimate these models was to *assign each citation and complaint to a facility-survey*, so that we could construct per-resident rates and include characteristics reported in the Annual Survey as covariates in the models.

For the deficiency citations data, this process was straightforward. Since facility inspections and Annual Surveys are both conducted during the facility's yearly recertification process, facility characteristics (recorded

in the Annual Survey) can be easily matched to data on deficiency citations using the "certification date" field. Therefore, we matched each citation to the facility-survey from the same facility (represented by the unique facility-level identifier) carrying the same certification date.

Unlike deficiency citations, however, complaints can be submitted throughout the year, and thus cannot be matched to a particular Annual Survey with the certification date alone. Our aim was to match a given complaint to the facility-survey whose reported facility characteristics best represented the facility's characteristics on the date the complaint was filed. To accomplish this goal, we relied on facility-survey certification dates *and* complaint submission dates. Specifically, for any given facility, we carried out the following procedure:[16]

1) **Since Annual Surveys are spaced approximately one year apart, we removed all complaints that were submitted *more than six months* before the facility's very first, or after its very last, facility-survey.** We do not know whether complaints filed more than six months before the facility's very first facility-survey, or after its very last facility-survey, in our dataset correspond well to the characteristics reported in the closest recorded Annual Survey. To take an extreme (and in our context, implausible) example, one might worry that if we match a complaint filed in January of 2019 to a facility-survey from May of 2013, the facility-level characteristics reported in the last Annual Survey (from May of 2013) might not correspond well to the characteristics of the facility on the date the complaint was filed.

   Importantly, this procedure did not remove any facility-surveys from the dataset; it merely excluded a relatively small number of complaints (3,641 complaints, or 7% of all complaints submitted from 2008 through 2017) from the complaints models.

2) **We assigned complaints submitted six months or less before the facility's very first certification date to the first facility-survey. Analogously, we assigned complaints submitted six months or less *after* the facility's last certification date to the last facility-survey**. Because Annual Surveys are spaced approximately one year apart, we assumed that the characteristics of the first (last) facility-survey were reasonably representative of complaints filed up to six months before (after) that facility-survey.

3) **We assigned complaints submitted between two facility-surveys to the closer of the two**. For example, if Annual Surveys were conducted on January 1, 2010 and January 1, 2011, a complaint filed anytime in May 2010 would be assigned to the January 1, 2010 facility-survey, whereas a complaint filed anytime in August 2010 would be assigned to the January 1, 2011 facility-survey. If a complaint was exactly between two facility-surveys, we randomly assigned it to one of the two facility-surveys.

All of the complaints models that we present in the paper were estimated by assigning complaints to facility-surveys using this strategy. As we discuss further in Section 2.3, as a robustness check, we estimated an *alternative* set of complaints models in which we "trimmed" (i.e., removed) the very first, and the very last, facility-surveys for all facilities, as well as their corresponding complaints, to account for the possibility that some of these complaints might in fact have more closely reflected the facility characteristics reported in facility-surveys that took place less than twelve months before the first, or after the last, facility-survey. In short, the "trimmed" dataset used for this robustness check *only* includes those complaints that we could confidently assign to their nearest facility-surveys. However, this approach came at the price of substantially reducing the number of facility-surveys in our dataset. The results of this robustness check are presented in the Results Packet. (Note that we only run our models on a dataset that has been both cleaned *and* trimmed, which is why our Results Packet refers to our trimmed dataset as the "cleaned, trimmed" dataset. We outline how we order the cleaning and trimming procedures in our code on page 24 below.)

---

[16] We match complaints with facility-surveys in "1_Merge/Code/Merge.R."

III) Construction of the Complaint-Days Variable

As discussed above, Annual Surveys are not always spaced 12 months apart. Failing to control in some fashion for the *time elapsed between adjacent surveys* in our complaints models could inadvertently bias our results, since the more months a facility is in operation, the more clients it is likely to have, the more care employees, providers, family members, and other stakeholders it is likely to interact with, etc.—and thus *ceteris paribus*, the more complaints one might expect to be filed against it. To take an extreme (and once again, in our context, implausible) example, an entity against which four complaints are filed in ten years might be expected to provide a *higher* level of care than a facility against which three complaints are filed in only two months. Yet a simple count model does not automatically account for the difference in time periods.

To account for the fact that complaints may be partly a function of the time elapsed between adjacent Annual Surveys, we created a variable called "complaint-days," which records the number of days for which complaints can be assigned to a given facility-survey. We include complaint-days as a covariate in our OLS models, and as an exposure term[17] in our negative binomial models, in which total complaints or substantiated complaints is the dependent variable.

We construct the complaint-days variable[18] using the following three-step procedure for each facility-survey:

1) **First, we determined the facility-survey's "prior range"—that is, the maximum number of days that a complaint could be submitted *before* the facility-survey and still be assigned to it in our assignment procedure.**
   a. **If a given facility-survey was not the facility's first, its prior range is defined as half the number of elapsed days between it and the previous facility-survey**. Using the facility in Table A11 as an example, there were 336 days between the 03/05/2013 and 02/04/2014 facility-surveys. Thus, the prior range for the 02/04/2014 facility-survey was 168 days. (As we explain above, complaints submitted exactly on 08/20/2013, the midpoint between the two facility-surveys, would be randomly assigned to either the 03/05/2013 or the 02/04/2014 facility-survey.)
   b. **If the facility-survey was the facility's *first* observation, then its prior range was a half-year, or 183 days.** This is because, as we outline above, complaints submitted prior to the first facility-survey are assigned to it only if they were submitted a half-year or less before the first certification date.

2) **Secondly and analogously, we determined the facility-survey's "after range"—that is, the maximum number of days that a complaint could be submitted *after* the facility-survey and still be assigned to it in our assignment procedure**.
   a. **If the facility-survey was not the facility's last observation, its after range was defined as the half the number of days between it and the subsequent facility-survey**. For example, there were 366 days between the 01/27/2016 and 01/27/2017 facility-surveys (2016 was a leap year). Thus, the after range for the 01/27/2016 facility-survey is 183. (As we explain above, complaints submitted on 07/28/2016, the midpoint between the two facility-surveys, would be randomly assigned to either the 01/27/2016 or 01/27/2017 facility-survey).
   b. **If the facility-survey is the facility's *last* observation, then its after range is the *lesser of* a) 183 days, and b) the number of elapsed days between the last facility-survey and**

---

[17] An exposure term is a variable that can be included in negative-binomial regressions to effectively transform count data into rates, thereby controlling for the fact that different facility-surveys can encompass different periods of time.

[18] The complaint-days variable is created in "1_Merge/Code/Merge.R."

04/26/2017—the last date on which complaints were included in our dataset.[19] For example, if a facility-survey occurred on 01/27/2017, there would only be 87 elapsed days until 04/26/2017. Complaints submitted after 04/26/2017 are not included in our dataset, so the 01/27/2017 facility-survey can only be assigned to complaints submitted up until that date (a period of 87 days after the facility-survey). Note that if a facility's final facility-survey occurred more than 183 days before 04/26/2017, we would only consider complaints submitted within 183 days of the final facility-survey.

3) **Finally, we define a facility-survey's complaint-days (i.e., the length, in days, of the period in which a complaint could be assigned to the facility-survey) as the sum of its prior range and after range**.

Table A11: Facility with Prior Range, After Range, and Total Complaint-Days Exposure Values

Certification Date	Prior Range (in days)	After Range (in days)	Complaint-Days (total exposure)
01/27/2017	183	87	270
01/27/2016	170.5	183	353.5
02/20/2015	190.5	170.5	361
02/04/2014	168	190.5	358.5
03/05/2013	161.5	168	329.5
04/16/2012	181.5	161.5	343
04/19/2011	177.5	181.5	359
04/29/2010	186.5	177.5	364
04/21/2009	183	186.5	369.5

To complete the three tasks outlined in this section (assigning complaints to facility-surveys, constructing the trimmed dataset, and complaint-days), we used the following sequence. First, we flagged the first and last facility-survey for each facility, so that they could be trimmed from the dataset at a later stage; created the complaint-days exposure term; and assigned complaints to their respective facility-surveys. Secondly, we implemented the cleaning procedures for ownership type and resident count, widely-spaced annual surveys, and direct-care staff ratio (as explained in Section 1.2 (III), the last two procedures only affect the complaints and direct-care staff models, respectively). Finally, to create the trimmed dataset, we trimmed the (previously flagged) first and last facility-surveys that had not already been removed by the cleaning procedure. Sequencing the procedures in this manner prevented us from improperly dropping the last facility-survey for a given facility (while cleaning the ownership status variable) *and* the same facility's second-to-last facility survey (as part of the trimming process) from the trimmed dataset, which might occur if we flagged first and last facility-surveys *after* cleaning the ownership type variable. s

---

[19] Our data were extracted on May 1, 2017, but the last complaint in our dataset was filed on April 26, 2017. Because the data were extracted in the morning of May 1, we would not expect CMS to include complaints filed on that day. That said, it was not immediately clear whether the lack of complaints in the period between April 27 - April 30 suggested that there were *no* complaints filed in that period, or if our dataset simply did not include them due to lags in the CMS reporting system. For the purposes of constructing complaint-days, we therefore had to determine whether it was plausible for *no* complaints to be filed for the four consecutive days of April 27 - April 30. Throughout our dataset—which spans the period of October 1, 2004 through April 26, 2017—the largest number of consecutive days without complaints is four days, and lapses of this length occur only five times. In all five cases, the four-day period corresponded to the four days (Thursday, Friday, Saturday, and Sunday) of that year's Thanksgiving break. Because there were no major holidays in late April 2017, we do not think it is plausible for there to be no complaints filed in the period between April 27 and April 30. Therefore, although our data were extracted on May 1, we believe that April 26 was the last day a complaint could be filed and still be included in our dataset.

IIII) Description of Deficiency Citations and Complaints Variables

Table A12: Variables from ASPEN Database Not Reported in the Annual Survey

Variable	Description
Total Deficiency Citations	The number of total deficiency citations recorded in CMS' ASPEN system for a given facility-survey. Deficiency citations are matched to facility-surveys in the Annual Survey dataset through the certification date field.
Total Deficiency Citations per Resident	The number of total deficiency citations divided by the number of residents reported by the ICF-IID in field W33 of the Annual Survey.
Total Complaints	The number of complaints recorded in CMS' ASPEN system for a given facility-survey. Complaints are assigned to facility-surveys in the Annual Survey dataset in the procedure outlined in Section 1.3 (II).
Total Complaints per Resident	The number of total complaints divided by the number of residents reported by the ICF-IID in field W33 of the Annual Survey.
Substantiated Complaints	The number of total complaints that are marked as "Substantiated" (as opposed to "Unsubstantiated"). A complaint is marked as "Substantiated" if a complaint investigation determines that the ICF-IID in question was indeed deficient (with respect to standards set by CMS for ICF-IIDs that receive Medicaid funds).
Substantiated Complaints per Resident	The number of substantiated complaints divided by the number of residents reported by the ICF-IID in field W33 of the Annual Survey.
Complaint-Days	The number of days during which complaints could be assigned to a given facility-survey using the procedure outlined in Section 1.3 (II).

2) DESCRIPTION OF ALTERNATIVE DATASETS AND SPECIFICATIONS

In the paper, we only present results from models estimated on the *cleaned* version of the original (imbalanced) dataset (described in Section 1.2(III)), using the specifications described in the paper. To check the robustness of our findings, however, we also estimate the same models on the *uncleaned* dataset, and on three *alternative* datasets in which we use statistical procedures to improve covariate balance across ownership types. Finally, we perform an additional robustness check for the complaints models: we re-estimate them on a *trimmed* version of the cleaned dataset.

2.1) Cleaned vs. Uncleaned Dataset

As mentioned above, the models we present in the paper were estimated on the *cleaned* dataset. To check the robustness of our findings, we also estimated our models on the *uncleaned* dataset, on which we did not (with one exception) perform the various cleaning procedures described above.[20] (The one exception was the removal of facility-surveys reporting an "other" ownership type: since these facility-surveys could not be meaningfully compared to facility-surveys that reported government, for-profit, or nonprofit ownership types, we removed them from *all* uncleaned datasets.) The results obtained from models estimated on the uncleaned dataset are referenced briefly throughout the paper, and presented in detail in the Results Packet.

As was described in Section 1.2 (III), we constructed our cleaned dataset by dropping facility-surveys that report an erroneous number of total residents and by cleaning—and when cleaning was not possible, dropping—facility-surveys with implausible reported ownership types. These two procedures produced the "baseline" cleaned dataset that we used for most models presented in the paper. In two instances, we removed additional facility-surveys from the cleaned dataset before estimating specific models. First, if a facility-survey was separated from an adjacent facility-survey by a period of 24 months or more, then for the reasons explained in detail above we removed it from the cleaned dataset used in the *complaints models.* Second, if we judged a facility-survey's reported direct-care staff ratio to be implausible, then for the reasons explained above we removed it from the cleaned dataset used in the *direct-care staff ratio models*.

As a consequence of the above procedures, the complaints and direct-care staff ratio models were estimated on *slightly different respective versions of the cleaned dataset*; moreover, both of the latter versions of the cleaned dataset differed slightly from the version used in other models. For example, if a facility-survey reported a valid ownership type but an implausible direct-care staff ratio, it would be included in the complaints models but excluded from the direct-care staff ratio models. Conversely, a facility-survey that reported a valid ownership type, but occurred more than two years before (or after) an adjacent facility-survey, would be included in the direct-care staff ratio models but excluded from the complaints models. Moreover, *both* of the facility-surveys just described would be included in the version of the cleaned dataset used to estimate the other models (deficiency citations; nursing staff ratio; and use of drugs, seclusion and restraint to control behavior).

2.2) Alternative Datasets to Improve Covariate Balance

The only point we intended to make is that the *average* resident count of government-owned facilities (51.7) far exceeds that of for-profit and nonprofit ones (8.1 and 10.2, respectively)

---

[20] We run our regressions on the cleaned and uncleaned versions of the dataset in the do-files found in "12_Regressions/2_Generate_Tables."

Government-owned ICFs are typically much larger than for-profit and nonprofit facilities. In 2016, for example, the average number of residents in government facilities was 51.7, the average numbers in for-profit and nonprofit facilities were only 8.1 and 10.2, respectively. It is well-known that extreme imbalance in covariates can bias coefficient estimates. Methodological approaches toward reducing covariate imbalance generally encompass two different strategies: "pruning" the dataset by removing facility-surveys that contribute a great deal to the imbalance; or assigning different weights to different facility-survey such that covariate values are similar across groups. Pruning-type approaches, which were the first to be described in the literature, are designed to improve covariate balance when only two groups are being compared (that is, when the covariate of interest can take on two values). Yet at least one of the more recent, weighting-type approaches has been modified to accommodate more than two groups (i.e., situations in which the covariate of interest can take on three or more different values) (Sävje et al., 2020).

Since our covariate of interest—ownership type—can take on three potential values, we implement one pruning-type and two weighting-type procedures in an effort to improve covariate balance between ownership groups. These three balancing procedures are performed on the "baseline" cleaned dataset referenced in Section 2.1 that is used in most regression models (i.e., all except for the complaints and direct-care staff ratio models). We introduce these three modified datasets—one pruned dataset, and two weighted datasets—as "alternative datasets" in page 13 of the paper, and refer to them as such throughout. To check the robustness of our main results, we re-estimate our regression models on the three alternative datasets[21] and present the results in the Results Packet.

I) Pruning-Type Approach

While conventional pruning-type procedures are designed to achieve covariate balance between two groups, we are interested in comparing *three* groups: government-owned, for-profit, and nonprofit facilities. To circumvent this challenge, we implemented a well-known "pruning" procedure (King et al., 2017) in two stages: (1) first, we implemented the pruning procedure to achieve reasonable covariate balance between two groups, for-profits and nonprofits; and then (2) we implemented the pruning procedure a second time to achieve covariate balance between the latter (combined) dataset and the data from government-owned facilities.

II) Weighting-Type Approaches

As noted above, methodological literature has advanced to the point that weighting-type methods can now be used to achieve covariate balance among two or more groups (i.e., cases in which the covariate of interest takes on three values), as is the case in our study (Sävje et al., 2020). To assign balancing weights, the algorithm first matches observations (i.e., facility-surveys) from different ownership groups into "matched" groups based on similarity in the respective observations' covariate values. In our case, we matched facility-surveys based on similarity in the facility characteristics used as covariates in our models, such as number of residents and age/diagnosis breakdowns. Then, for each matched group, the algorithm generates a weight for each observation based on the relative prevalence of that ownership type within the group. For example, if a matched group includes mostly nonprofit facilities, each nonprofit facility-survey in that group is "down weighted" relative to other facility-surveys in the group.

Importantly, given the structure of our dataset—i.e., the fact there are multiple facility-surveys for each facility, multiple facilities in each state, and multiple states in each year—there are different ways to implement the matching procedure and, thus, to assign weights to different facility-survey. To test the robustness of our findings, we used two different weighting procedures:

---

[21] We run these regressions in "13_Generate_Tables/2_Generate_Tables/Code/Generate_Tables_new.do."

1. **First, we weighted each facility-survey based on the facility-level characteristics reported in its Annual Survey**: In other words, the algorithm matched facility-surveys across different ownership groups based on similarities in the facility-surveys' respective facility characteristics, including number of residents, reported in each facility-survey. We implemented this procedure in two different ways: by initially stratifying the data by state (i.e., only matching facility-surveys in the same state), and *without* initially stratifying the data by state (i.e., by matching facility-surveys to each other regardless of whether they were located in the same state).[22] Since the non-stratified version of the facility-survey weights achieved better covariate balance across ownership groups, we used those weights for the "alternative datasets" used in our robustness checks.

2. **Secondly, we weighted each facility-survey based on the *median* characteristics of the facility across *all* facility-surveys in the dataset:** In this approach, we tried matching facility-surveys across ownership groups based on *median* characteristics reported for a given facility across all facility-surveys. For example, if a facility had seven facility-surveys with the same ownership type, we matched those seven observations to others in the dataset based on the *median* covariate values (including number of residents) observed across all seven facility-surveys. Note that because our goal was to achieve balance between ownership groups, if a facility switched from one ownership type to another during our period of study, we created two different sets of median facility characteristics—one set for each ownership type—on which to match. For example, if a facility with four facility-surveys reported a nonprofit ownership status for the first two facility-surveys and for-profit status for the last two facility-surveys, we matched the first and last pair of facility-surveys, respectively, based on the median facility characteristics calculated using the values reported in each pair. Here again, we implemented two variations of the weighting procedure: one in which we initially stratified the data by state and one in which we did not.[23] Since the non-stratified procedure again produced better covariate balance, we used the non-stratified weights in our alternative datasets to check the robustness of our findings.

After performing all of the procedures described above, we obtained three "alternative" (balanced) datasets—one obtained using a two-step pruning-type procedure and two obtained using weighting-type procedures—which we used in performing robustness checks. The three datasets are referred to as "cleaned, pruned," "cleaned, weighted by facility-survey characteristics," and "cleaned, weighted by median facility characteristics" throughout the Results Packet. The results of those robustness checks are referenced briefly to throughout the paper, and are presented in detail in the Results Packet.

2.3) Untrimmed vs. Trimmed Version of the Cleaned Dataset (in the Complaints Regressions):

As is described above in Section 1.3 (II), for the complaints models only, we created a "trimmed" dataset by dropping the first and last facility-survey for each facility (if they were still in the dataset after the cleaning procedure). Recall that the "trimmed" dataset *only* included complaints for which we could confirm that the facility-survey in our dataset was the closest one to the date on which the complaint was filed.

Because this trimming procedure removed quite a few facility-surveys from our dataset (7581 facility-surveys, or 18% of our cleaned dataset), the paper present results from complaints models estimated on the *untrimmed* versions of the cleaned dataset. However, as a robustness check, we present the results from complaints models estimated on the trimmed versions in the Results Packet as well.

---

[22] We implement both of these procedures in "8_Matching/Code/Matching.R."

[23] We perform both of these procedures in "8_Matching/Code/Matching.R."

3) SUMMARY STATISTICS

3.1) Proportion of Total Residents by Ownership Type in Uncleaned Dataset

Figure A5: Proportion of Total Residents from 2009-2017 in the Untrimmed Version of the Uncleaned Dataset[24]
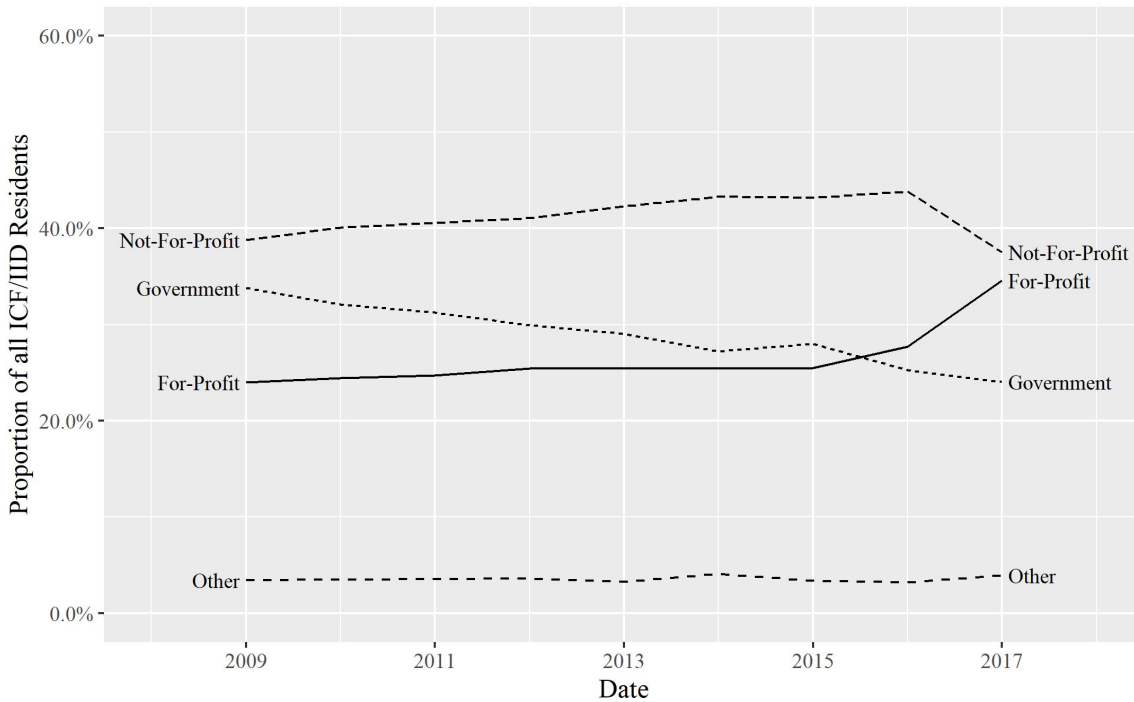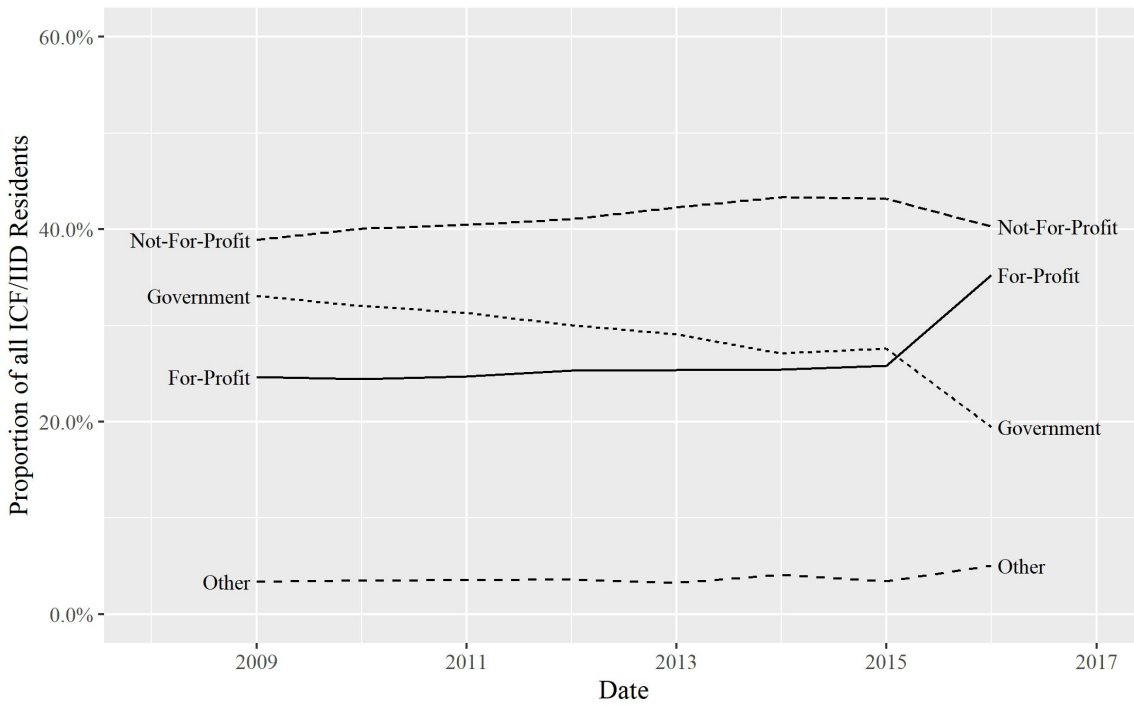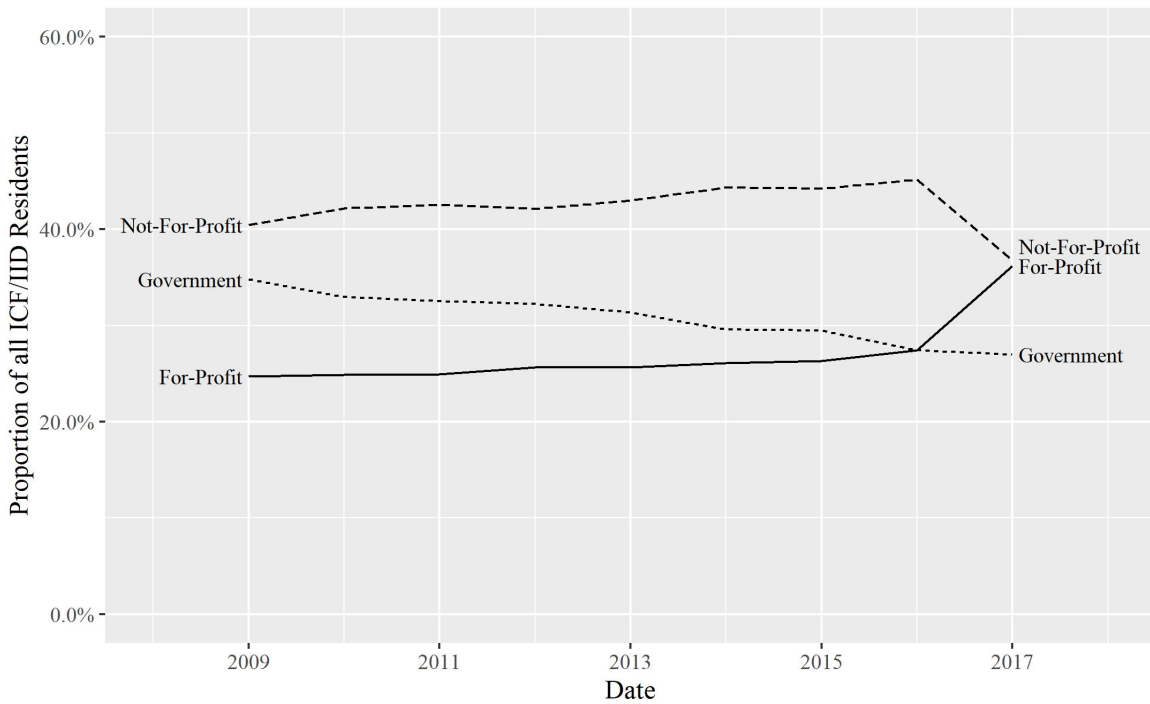

Figure A6: Proportion of Total Residents from 2009-2016 in the Trimmed Version of the Uncleaned Dataset

---

[24] Because we received our data in May 2017, our 2017 data does not contain Annual Survey data for all facilities, which helps explain the relatively stark changes in the ownership distribution between 2016 and 2017.

3.2) Proportion of Total Residents by Ownership Type in Cleaned Dataset

Figure A7: Proportion of Total Residents from 2009-2017 in the Untrimmed Version of the Cleaned Dataset[25]

Figure A8: Proportion of Total Residents from 2009-2016 in the Trimmed Version of the Cleaned Dataset

---

[25] Because we received our data in May 2017, our 2017 data does not contain Annual Survey data for all facilities, which helps explain the relatively stark changes in the ownership distribution between 2016 and 2017.
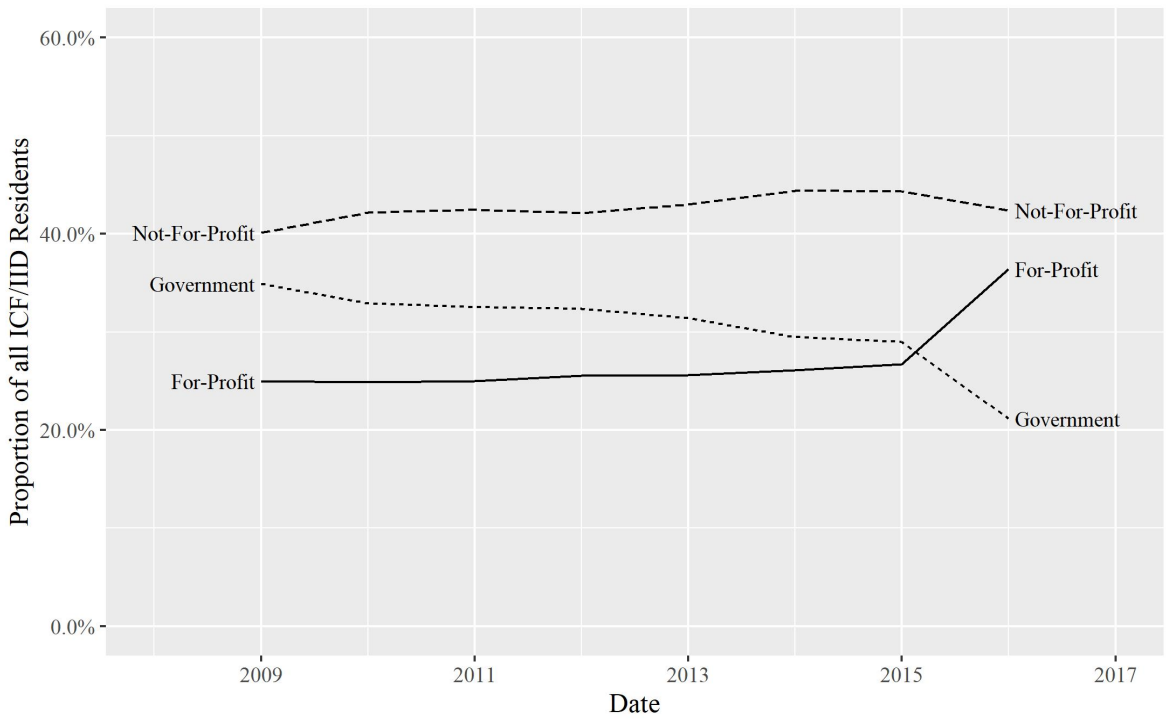
3.3) Facility Breakdown for Each Dataset by Ownership Type

Table A13: Facility Summary Statistics

Variable	Uncleaned			Cleaned, unadjusted			Cleaned, weighted by median facility characteristics			Cleaned, weighted by facility-year characteristics			Cleaned, pruned			Cleaned, unadjusted, trimmed		
	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.
Total number of facilities	2,942	3,734	459	2,258	2,976	349	2,258	2,976	349	2,258	2,976	349	2,186	2,846	325	2,239	2,926	344
Total number of facility-years	18,196	23,632	2,715	16,606	21,983	2,418	16,606	21,983	2,418	16,606	21,983	2,418	13,380	17,526	2,094	13,525	17,966	1,935
Total number of residents	152,624	249,111	177,353	139,529	231,408	169,015	139,529	231,408	169,015	139,529	231,408	169,015	114,509	189,706	80,580	115,129	191,304	140,648

3.4) Summary Statistics for Dependent Variables

Table A14: Dependent Variable Summary Statistics

Variable	Uncleaned			Cleaned, unadjusted			Cleaned, weighted by median facility characteristics			Cleaned, weighted by facility-year characteristics			Cleaned, pruned			Cleaned, unadjusted, trimmed		
	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.
Outcome variables issued by surveyors during ICF-IID (re)certification process																		
Total deficiency citations	3.88 (4.81)	3.19 (4.06)	4.89 (6.69)	3.88 (4.85)	3.15 (3.99)	5.03 (6.92)	4.19 (5.14)	3.26 (4.14)	3.52 (4.64)	4.18 (5.09)	3.26 (4.14)	3.61 (4.84)	3.86 (4.78)	3.19 (3.99)	4.30 (6.00)			
Total deficiency citations per resident	0.62 (0.84)	0.44 (0.62)	0.27 (0.47)	0.61 (0.84)	0.43 (0.61)	0.25 (0.44)	0.6 (0.85)	0.45 (0.64)	0.44 (0.62)	0.58 (0.82)	0.45 (0.63)	0.45 (0.64)	0.60 (0.79)	0.42 (0.59)	0.28 (0.46)			
Outcome variables reported by stakeholders																		
Total complaints	0.90 (2.47)	0.61 (2.53)	4.14 (22.38)	0.91 (2.48)	0.59 (2.35)	4.07 (16.48)	1.18 (4.22)	0.70 (3.20)	0.87 (5.58)	1.16 (3.85)	0.67 (2.66)	0.75 (5.17)	0.88 (2.51)	0.58 (2.39)	2.21 (11.52)	0.92 (2.51)	0.59 (2.42)	4.10 (15.86)
Total complaints per resident	0.13 (0.35)	0.07 (0.23)	0.05 (0.18)	0.13 (0.35)	0.07 (0.22)	0.05 (0.17)	0.13 (0.34)	0.07 (0.23)	0.05 (0.21)	0.13 (0.33)	0.07 (0.23)	0.05 (0.21)	0.13 (0.35)	0.06 (0.21)	0.04 (0.18)	0.13 (0.36)	0.07 (0.22)	0.05 (0.17)
Substantiated complaints	0.47 (1.62)	0.29 (1.29)	1.45 (11.33)	0.48 (1.65)	0.28 (1.26)	1.26 (5.97)	0.58 (2.53)	0.34 (1.67)	0.30 (2.19)	0.55 (2.06)	0.33 (1.45)	0.23 (1.94)	0.46 (1.69)	0.27 (1.29)	0.72 (4.26)	0.48 (1.65)	0.28 (1.28)	1.28 (5.9)
Substantiated complaints per resident	0.07 (0.23)	0.04 (0.15)	0.02 (0.08)	0.07 (0.24)	0.03 (0.14)	0.02 (0.07)	0.07 (0.22)	0.04 (0.15)	0.02 (0.08)	0.07 (0.22)	0.04 (0.15)	0.01 (0.07)	0.07 (0.23)	0.03 (0.13)	0.02 (0.07)	0.07 (0.24)	0.03 (0.14)	0.02 (0.07)
Outcome variables reported by ICF-IID personnel during (re)certification process																		
Direct care staff per resident	1.35 (4.22)	1.54 (3.32)	1.76 (1.22)	1.20 (0.51)	1.38 (0.55)	1.67 (0.65)	1.22 (0.53)	1.38 (0.55)	1.65 (0.67)	1.22 (0.53)	1.37 (0.54)	1.59 (0.65)	1.18 (0.5)	1.34 (0.53)	1.68 (0.67)			
Registered nurses per resident	0.06 (0.24)	0.08 (0.2)	0.16 (0.36)	0.06 (0.25)	0.08 (0.18)	0.17 (0.38)	0.06 (0.26)	0.08 (0.18)	0.21 (0.5)	0.06 (0.27)	0.08 (0.18)	0.19 (0.39)	0.05 (0.17)	0.07 (0.16)	0.18 (0.39)			
Use of drugs to control behavior	0.84 (0.37)	0.88 (0.32)	0.82 (0.38)	0.83 (0.37)	0.89 (0.32)	0.82 (0.39)	0.84 (0.37)	0.87 (0.33)	0.82 (0.38)	0.84 (0.37)	0.88 (0.33)	0.82 (0.39)	0.85 (0.36)	0.89 (0.31)	0.81 (0.39)			
Use of physical restraints	0.08 (0.28)	0.20 (0.40)	0.42 (0.49)	0.08 (0.28)	0.21 (0.41)	0.45 (0.50)	0.12 (0.32)	0.20 (0.40)	0.33 (0.47)	0.11 (0.31)	0.20 (0.40)	0.32 (0.47)	0.08 (0.27)	0.20 (0.40)	0.40 (0.49)			
Use of time out room	0.00 (0.06)	0.02 (0.14)	0.05 (0.23)	0.00 (0.06)	0.02 (0.15)	0.06 (0.23)	0.00 (0.06)	0.02 (0.14)	0.01 (0.11)	0.01 (0.08)	0.02 (0.14)	0.01 (0.11)	0.00 (0.06)	0.02 (0.14)	0.03 (0.17)			

Note. Unweighted means are presented for the four unweighted datasets ("Uncleaned," "Cleaned only," "Cleaned, pruned," and "Cleaned, unadjusted, trimmed"), with standard errors in parentheses. *Weighted* means are presented for the two weighted datasets ("Cleaned, weighted by median facility characteristics," and "Cleaned, weighted by facility-year characteristics), with weighted standard errors in parentheses.

3.5) Summary Statistics for Independent Variables

Table A15: Independent Variable Summary Statistics

Variable	Uncleaned			Cleaned, unadjusted			Cleaned, weighted by median facility characteristics			Cleaned, weighted by facility-year characteristics			Cleaned, pruned			Cleaned, unadjusted, trimmed		
	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.	For-Profit	Non-Profit	Gov.
Complaint-Days	363.19 (30.41)	360.91 (31.34)	364.69 (34.77)	363.36 (30.54)	360.99 (31.02)	365.16 (35.47)	363.43 (32.37)	361.53 (32.44)	366.59 (34.34)	363.55 (31.6)	361.27 (31.61)	366.19 (36.59)	363.38 (30.01)	361.09 (30.93)	365.34 (35.88)	366.58 (26.46)	363.32 (28.38)	367.52 (34.49)
Number of residents	8.39 (13.88)	10.54 (16.34)	65.32 (106.2)	8.4 (13.65)	10.53 (15.36)	69.9 (110.66)	11.92 (24.76)	12.26 (24.47)	14.47 (35.31)	12.11 (25.83)	12.2 (24.42)	13.86 (34.44)	8.56 (13.87)	10.82 (15.59)	38.48 (54.74)	8.51 (14.1)	10.65 (15.56)	72.69 (112.5)
1-6 residents	0.75 (0.43)	0.53 (0.5)	0.25 (0.43)	0.75 (0.43)	0.53 (0.5)	0.23 (0.42)	0.66 (0.47)	0.56 (0.5)	0.39 (0.49)	0.64 (0.48)	0.57 (0.49)	0.39 (0.49)	0.73 (0.45)	0.5 (0.5)	0.25 (0.43)	0.74 (0.44)	0.52 (0.5)	0.21 (0.4)
7-15 residents	0.2 (0.4)	0.37 (0.48)	0.27 (0.45)	0.2 (0.4)	0.37 (0.48)	0.27 (0.44)	0.23 (0.42)	0.33 (0.47)	0.47 (0.5)	0.25 (0.44)	0.32 (0.47)	0.49 (0.5)	0.22 (0.41)	0.39 (0.49)	0.3 (0.46)	0.21 (0.41)	0.38 (0.48)	0.27 (0.45)
16-34 residents	0.02 (0.15)	0.06 (0.24)	0.13 (0.34)	0.02 (0.15)	0.07 (0.25)	0.13 (0.34)	0.06 (0.23)	0.05 (0.22)	0.09 (0.28)	0.05 (0.23)	0.05 (0.22)	0.06 (0.24)	0.03 (0.16)	0.07 (0.26)	0.15 (0.36)	0.02 (0.15)	0.07 (0.25)	0.13 (0.34)
35 residents or more	0.02 (0.15)	0.04 (0.19)	0.35 (0.48)	0.02 (0.15)	0.04 (0.19)	0.37 (0.48)	0.05 (0.22)	0.05 (0.22)	0.05 (0.22)	0.05 (0.22)	0.05 (0.22)	0.05 (0.22)	0.02 (0.15)	0.04 (0.19)	0.29 (0.46)	0.02 (0.16)	0.04 (0.19)	0.38 (0.49)
Proportion under 22	0.07 (0.19)	0.09 (0.23)	0.04 (0.1)	0.07 (0.18)	0.09 (0.23)	0.04 (0.1)	0.08 (0.2)	0.08 (0.22)	0.04 (0.13)	0.08 (0.2)	0.08 (0.22)	0.04 (0.13)	0.04 (0.1)	0.04 (0.12)	0.03 (0.1)	0.07 (0.18)	0.09 (0.23)	0.04 (0.1)
Proportion over 65	0.09 (0.14)	0.09 (0.15)	0.12 (0.16)	0.09 (0.14)	0.09 (0.15)	0.12 (0.16)	0.09 (0.14)	0.09 (0.15)	0.11 (0.17)	0.09 (0.14)	0.09 (0.14)	0.09 (0.15)	0.08 (0.12)	0.08 (0.13)	0.12 (0.16)	0.08 (0.14)	0.09 (0.14)	0.12 (0.15)
Proportion male	0.57 (0.37)	0.57 (0.34)	0.62 (0.29)	0.57 (0.37)	0.57 (0.34)	0.63 (0.28)	0.57 (0.35)	0.57 (0.35)	0.62 (0.34)	0.57 (0.35)	0.57 (0.35)	0.62 (0.35)	0.58 (0.37)	0.58 (0.34)	0.63 (0.29)	0.56 (0.37)	0.57 (0.34)	0.63 (0.27)
Proportion severe/ profound ID	0.51 (0.34)	0.56 (0.33)	0.7 (0.29)	0.5 (0.34)	0.56 (0.33)	0.72 (0.28)	0.53 (0.33)	0.55 (0.33)	0.62 (0.34)	0.54 (0.33)	0.55 (0.33)	0.6 (0.34)	0.5 (0.34)	0.58 (0.32)	0.71 (0.28)	0.51 (0.34)	0.57 (0.33)	0.72 (0.27)
Proportion autistic	0.11 (0.18)	0.15 (0.23)	0.14 (0.17)	0.11 (0.17)	0.15 (0.23)	0.14 (0.17)	0.12 (0.19)	0.14 (0.22)	0.14 (0.19)	0.12 (0.19)	0.14 (0.21)	0.14 (0.19)	0.09 (0.15)	0.13 (0.19)	0.13 (0.17)	0.1 (0.17)	0.15 (0.23)	0.14 (0.17)
Proportion cerebral palsy	0.21 (0.25)	0.21 (0.24)	0.19 (0.21)	0.21 (0.25)	0.22 (0.24)	0.19 (0.21)	0.21 (0.24)	0.22 (0.25)	0.16 (0.2)	0.21 (0.24)	0.22 (0.24)	0.17 (0.21)	0.18 (0.22)	0.2 (0.22)	0.19 (0.22)	0.21 (0.25)	0.21 (0.24)	0.2 (0.21)
Proportion epilepsy	0.35 (0.26)	0.35 (0.25)	0.44 (0.24)	0.35 (0.26)	0.36 (0.25)	0.45 (0.24)	0.36 (0.26)	0.35 (0.26)	0.4 (0.25)	0.36 (0.25)	0.36 (0.26)	0.37 (0.25)	0.34 (0.25)	0.35 (0.24)	0.45 (0.25)	0.35 (0.26)	0.35 (0.25)	0.45 (0.24)
Proportion language impairment	0.44 (0.39)	0.49 (0.39)	0.62 (0.37)	0.44 (0.39)	0.5 (0.39)	0.64 (0.36)	0.46 (0.38)	0.49 (0.38)	0.54 (0.39)	0.47 (0.38)	0.49 (0.39)	0.5 (0.39)	0.43 (0.38)	0.49 (0.38)	0.63 (0.37)	0.44 (0.39)	0.49 (0.38)	0.64 (0.36)
Proportion hearing impairment	0.13 (0.18)	0.14 (0.17)	0.16 (0.18)	0.13 (0.18)	0.14 (0.17)	0.17 (0.18)	0.13 (0.18)	0.14 (0.18)	0.13 (0.17)	0.13 (0.18)	0.14 (0.18)	0.13 (0.15)	0.11 (0.14)	0.12 (0.14)	0.15 (0.16)	0.13 (0.18)	0.14 (0.17)	0.17 (0.18)

34

Proportion visual impairment	0.45 (0.3)	0.43 (0.29)	0.42 (0.27)	0.45 (0.3)	0.43 (0.29)	0.43 (0.26)	0.44 (0.29)	0.44 (0.29)	0.42 (0.29)	0.44 (0.29)	0.43 (0.29)	0.42 (0.28)	0.45 (0.29)	0.43 (0.28)	0.42 (0.26)	0.45 (0.3)	0.43 (0.29)	0.43 (0.26)
Proportion nonambulatory	0.56 (0.45)	0.55 (0.44)	0.6 (0.41)	0.55 (0.45)	0.55 (0.44)	0.61 (0.4)	0.56 (0.44)	0.56 (0.44)	0.6 (0.44)	0.56 (0.45)	0.55 (0.44)	0.57 (0.44)	0.54 (0.45)	0.53 (0.44)	0.63 (0.41)	0.55 (0.45)	0.55 (0.44)	0.61 (0.4)
Proportion with medical care plan	0.2 (0.39)	0.18 (0.37)	0.28 (0.43)	0.2 (0.39)	0.17 (0.36)	0.29 (0.44)	0.19 (0.39)	0.19 (0.38)	0.21 (0.39)	0.19 (0.38)	0.19 (0.38)	0.19 (0.38)	0.17 (0.37)	0.14 (0.34)	0.28 (0.43)	0.19 (0.39)	0.17 (0.36)	0.29 (0.43)

Note. *Unweighted* means are presented for the four unweighted datasets ("Uncleaned," "Cleaned only," "Cleaned, pruned," and "Cleaned, unadjusted, trimmed"), with standard errors in parentheses. *Weighted* means are presented for the two weighted

datasets ("Cleaned, weighted by median facility characteristics," and "Cleaned, weighted by facility-year characteristics), with weighted standard errors in parentheses.

3.6) How We Calculate the Descriptive Statistics in the Manuscript

1) **Average number of residents per facility by ownership type:**
   a. <u>Use in paper</u>: "While ICF-IIDs vary in size and in ownership type, government-owned facilities are generally larger, with an average of 65.3 residents per facility, as compared to average resident counts of 8.4 and 10.5, respectively, among for-profit and nonprofit facilities."
   b. <u>How we calculated the numbers</u>: For a given ownership type (i.e., government-owned, for-profit, or nonprofit), we calculate the average number of residents across all facility-surveys of that ownership type in our uncleaned dataset (CMS, 2017).

Total number of ICF-IIDs in 2009 and 2016:
   c. <u>Use in paper</u>: "Between 2009 and 2016, for example, the total number of ICF-IIDs in the U.S. fell from 5,457 to 5,003."
   d. <u>How we calculated the numbers</u>: We use CMS's publicly-available provider lists for 2009 and 2016 (CMS, 2009, 2016). These two data sets list 5,457 and 5,0003 unique ICF-IIDs, respectively.

2) **Average number of residents per facility in 2009 and 2016:**
   a. <u>Use in paper</u>: "[T]he average number of residents per [ICF-IID] facility decreased from 13.7 [in 2009] to 11.6 [in 2016]."
   b. <u>How we calculated the numbers</u>: We calculate the average number of residents across *all* 2009 and 2016 facility-surveys, respectively, in our uncleaned dataset (CMS, 2017). Note that because of the erratic timing of Annual Surveys, a small number of facilities are surveyed twice in 2009 and/or 2016 (and are thus included twice, once per survey, in the calculation of their respective ownership groups' means'), while others are not surveyed in 2009 and/or 2016 at all.

3) **Percentage of ICF-IID residents by facility ownership type in 2009 and 2016:**
   a. <u>Use in paper</u>: "Between 2009 and 2016, the percentage of ICF-IID residents living in government-owned facilities fell from 33.8% to 25.2%. During the same period, the share of ICF-IID residents living in nonprofits rose from 38.8% to 43.9% and the percentage residing in for-profits rose from 24.0% to 27.7%."
   b. <u>How we calculated the numbers</u>: We calculate the proportion of total ICF-IID residents living in facilities of a given ownership type in 2009 and 2016, respectively, by (1) summing the total number of residents across all facility-surveys of that ownership type in that year and (2) dividing that value by the sum of residents across *all* facility-surveys in that year. We calculate these values using the raw, uncleaned dataset (CMS, 2017). Because this version of the data set includes a small number of facilities whose ownership type is reported as "other," the percentages listed in the paper do not add up to 100%. (The uncleaned dataset used for the regressions presented in the paper has these "other" facilities removed.) As with the previous statistics, a small number of facilities are surveyed twice in 2009 and/or 2016 (and are thus included twice, once per survey, in the calculations), while others are not surveyed in 2009 and/or 2016 at all.

4) **Percent of individuals with I/DD receiving LTSS, but not living with family members, who live in ICF-IIDs specifically:**
   a. <u>Use in paper</u>: "First, as of this writing, approximately 75,000 individuals—roughly 14% of all individuals with I/DD who receive LTSS from state agencies but do not live with family members—reside in ICF-IIDs nationwide."

b. <u>How we calculated the numbers</u>: Larson et al. (2020) estimate that in 2017, 73,855 people lived in ICF-IIDs across the country, while there were 516,505 people receiving LTSS from state agencies across the country that did not live with family members.

References

42 C.F.R. § 483.430.

42 C.F.R. § 483.450.

Centers for Medicare & Medicaid Services. (2009). *2009 POS [Provider of Service] File* [Data set]. https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services/POS2009

Centers for Medicare & Medicaid Services. (2013a). *Intermediate care facilities for individuals with intellectual disabilities survey report: Form-3070G*. Retrieved May 3, 2021, from https://www.cms.gov/Medicare/CMS-Forms/CMS-Forms/Downloads/CMS3070G.pdf

Centers for Medicare & Medicaid Services. (2013b, December 6). *State Operations Manual (SOM) Appendix J, guidance to surveyors: intermediate care facilities for individuals with intellectual disabilities, nomenclature revisions* (Pub. 100-07 State Operations Provider Certification, Transmittal 94). https://www.cms.gov/Regulations-and-Guidance/Guidance/Transmittals/Downloads/R94SOMA.pdf

Centers for Medicare & Medicaid Services. (2016). *2016 POS [Provider of Service] File* [Data set]. https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services/POS2016

Centers for Medicare & Medicaid Services. (2017). [Unpublished raw data on Annual Surveys of intermediate care facilities for individuals with intellectual disabilities (ICFs); deficiency citations given to ICFs; and complaints filed against ICFs]. Quality Improvement Evaluation System (QIES).

Centers for Medicare & Medicaid Services. (2018, April 13). *State Operations Manual: Appendix J*. Retrieved May 3, 2021, from https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/downloads/som107ap_j_intermcare.pdf

Centers for Medicare & Medicaid Services. (2019, July 19). *State Operations Manual: Chapter 5 – Complaint procedures*. Retrieved May 3, 2021, from https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Downloads/som107c05pdf.pdf

Centers for Medicare & Medicaid Services. (2021, April 16). *Provider of Service Current Files.* Retrieved May 3, 2021, from https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services

Karon, S.L., & Beutel, P. (2001). *Intermediate Care Facilities for the Mentally Retarded: Facility and client characteristics, 1999*. University of Wisconsin-Madison, Center for Health Systems Research and Analysis. https://www.semanticscholar.org/paper/Intermediate-Care-Facilities-for-the-Mentally-and-Karon-Beutel/c0bf3dcfc0eb84a96216177e41d10e900e7532c6#paper-header

Larson, S. A., Eschenbacher, H. J., Taylor, B., Pettingell, S., Sowers, M., & Bourne, M. L. (2020). *In-home and residential long-term supports and services for persons with intellectual or developmental disabilities: Status and trends through 2017*. University of Minnesota Institute on Community Integration. https://ici-s.umn.edu/files/aCHyYaFjMi/risp_2017

King, G., Lucas, C., & Nielsen, R. A. (2017). The balance-sample size frontier in matching methods for causal inference. *American Journal of Political Science*, *61*(2), 473–489. https://doi.org/10.1111/ajps.12272

Sävje, F., Higgins, M. J., & Sekhon, J. S. (2020). Generalized full matching. *Political Analysis.* Advance online publication. https://doi.org/10.1017/pan.2020.32

INTERMEDIATE CARE FACILITIES FOR INDIVIDUALS WITH INTELLECTUAL DISABILITIES
SURVEY REPORT

1. Name of Facility

2. Street Address	3. City and/or County	4. State	5. ZIP Code

6. Medicaid Provider Number	7. Name of CEO	8. Telephone No.

9. State/Region Code	10. State/County Code	11. Dates of Survey (mm/dd/yyyy)
W2	W3	Begin: _____ W4 End: _____ W5

12. Type of Ownership or Control (enter number in box below)
- 1. Private (non-profit)
- 2. Private (proprietary)
- 3. State
- 4. City/Town
- 5. County
- 6. City/County
- 7. Other (specify): _____   W6

13. Is this ICF/IID a distinct part of a Hospital, SNF or NF? (check one)
☐ Yes ☐ No   W7

14. If "Yes" to block 13, indicate either:
A. Hospital Provider Number:
B. SNF Provider Number:
C. NF Provider Number: W8

15. Survey Team Composition
Column 1: Indicate the number of disciplines represented on the Survey team.
Column 2: Of the number in Column 1 represented on the Survey team, indicate the number who also qualify as a QIDP. Indicate Name(s) and Title(s) on last page of this form.

W9   W10

	Col 1	Col 2
A. Administrator		
B. Nurse		
C. Dietitian		
D. Pharmacist		
E. Records Administrator		
F. Social Worker		
G. LSC Specialist		
H. Laboratorian		
I. Sanitarian		
J. Therapist		
K. Physician		
L. Psychologist		
M. Other (specify): _____		
N. Total number of Surveyors onsite W11		
O. Total number of QIDP Surveyors onsite W12		

16. Facility Data

A. Is this ICF/IID a residential unit within a larger organization or agency in the State that provides residential services to individuals with intellectual disabilities? (check one)
☐ Yes ☐ No If "No", proceed to item C.   W13

B. If "Yes," indicate name and address of larger organization.

Name: _____

Address: _____

City: _____ State: _____ Zip Code: _____

Name of CEO: _____

Total Number of Beds: ☐☐☐ W14   Total Number of Clients: ☐☐☐ W15
(including ICF/IID clients directly served)

C. Total Number of ICF/IID Clients: ☐☐☐   W16

D. Is this ICF/IID community-based? *(check one)* ☐ Yes ☐ No   W17

E. Total number of ICF/IID beds under this Provider No: ☐☐☐   W18

F. Total number of discrete living units under this Provider No: ☐☐☐   W19

G. Age range of clients served: from ☐☐ W20 to ☐☐☐ W21

H. Total number of off-campus day program sites used by ICF/IID clients: ☐☐   W22

17. Staffing: List the full time equivalents who function in this capacity:

A. Direct Care Personnel (483.430(d)(3))	☐☐☐☐.☐☐	W23
B. Registered Nurse (483.480(d)(3))	☐☐☐☐.☐☐	W24
C. Licensed Voc./Practical Nurse (483.480(d)(2))	☐☐☐☐.☐☐	W25
D. Total Personnel *(List the Full Time Equivalent for all employees)*	☐☐☐☐.☐☐	W26

18. Off-Campus Day Programs:

A. How many clients in the sample attend off-campus day programs? ☐☐☐   W27

B. In how many off-campus day program sites was an observation done by the Surveyor? ☐☐☐   W28

FORM CMS-3070G (03/13)   4

19. Individual Characteristics
(**NOTE**: The total number in Items B–L (Col.(a)) may exceed the facility's population because some clients have multiple disabilities)

A. AGE AND SEX

(1) Age

under 22(a)	W29
22-45 (b)	W30
46-65 (c)	W31
66+ (d)	W32
Total:	W33

(2) Sex

Male	W34
Female	W35
Total:	W36

B. DISABILITIES

(1) Intellectual Disability

Mild	W37
Moderate	W38
Severe	W39
Profound	W40
Total:	W41

(2) Autism	W42
(3) Cerebral Palsy	W43
(4) Epilepsy	
Controlled	W44
Uncontrolled	W45
Total:	W46

C. OTHER DISABILITIES

(1) Non-ambulatory

Mobile	W47
Non-Mobile	W48
Total:	W49

(2) Speech/Language Impairment	W50
(3) Hearing Impairment	
Hard of Hearing	W51
Deaf	W52
Total:	W53

(4) Visual Impairment

Impaired	W54
Blind	W55
Total:	W56

D. MEDICAL CARE PLAN	W57
E. DRUGS TO CONTROL BEHAVIOR	W58
F. PHYSICAL RESTRAINTS	W59
G. TIME-OUT ROOMS	W60
H. APPLICATION OF PAINFUL OR NOXIOUS STIMULI	W61
I. NUMBER ATTENDING OFF-CAMPUS DAY PROGRAMS	W62
J. NUMBER OF COURT ORDERED ADMISSIONS	W63
K. NUMBER OF CLIENTS OVER AGE 18 WITH A LEGAL GUARDIAN ASSIGNED BY THE COURT	W64

L. OTHER (specify)

(1) _____	W65
(2) _____	W66
(3) _____	W67

INTERMEDIATE CARE FACILITIES FOR INDIVIDUALS WITH INTELLECTUAL DISABILITIES
SURVEY REPORT

M. ALLEGATIONS OF ABUSE AND NEGLECT

No. of allegations of abuse investigated (a)	W68
No. of allegations of neglect investigated (b)	W69
Total:	W70

N. NUMBER OF DEATHS

No. of deaths related to unusual incidents (a)	W71
No. of deaths related to restraints (b)	W72
No. of deaths for any reason (c)	W73
Total:	W74

ALLEGATIONS OF ABUSE AND NEGLECT AND NUMBER OF DEATHS
DATA ENTRY INSTRUCTIONS

M. ALLEGATION OF ABUSE AND NEGLECT

(W68) Number of allegations of abuse investigated.
(W69) Number of allegation of neglect investigated.

According to 42CFR §488.301:
Abuse is the willful infliction of injury, unreasonable confinement, intimidation or punishment with resulting physical harm, pain or mental anguish.

Neglect is the failure to provide goods and services necessary to avoid physical harm, mental anguish or mental illness.

Consistent with the referenced definitions, enter the number of allegations of abuse and or neglect investigated, including investigations resulting from complaints, follow ups, initials or recertifications.
If there is no information to report, leave the field blank.

(W70) Total
This field represents a combined total of W68 (allegations of abuse investigated) and W69 (allegations of neglect investigated). The total for this field is program generated therefore, no data input is necessary.

N. NUMBER OF DEATHS

(W71) Number of deaths related to unusual incidents.
Insert the number of deaths that occurred as a result of unusual incidents. This includes all unexpected or unanticipated deaths not included in W72 or W73.

(W72) Number of death related to restraints.
Insert the number of deaths that occurred as a result of the use of restraints.

(W73) Number of deaths for any reason.
Insert the number of deaths occurring for any reason. Do not include information contained is W71 and W72 above.

(W74) Total
This field represents a combined total of W71 (number of deaths related to unusual incidents), W72 (number of deaths related to restraints), and W73 (number of deaths for any reason).

The total for this field is program generated; therefore, no data input is necessary.

FORM CMS-3070G (03/13)   4