Does Quality Disclosure Improve Quality? Responses to the Introduction of Nursing Home Report Cards in Germany

Annika Herr, Thu-Van Nguyen, Hendrik Schmitz

February 2015
Does Quality Disclosure Improve Quality? Responses to the Introduction of Nursing Home Report Cards in Germany

Annika Herr† Thu-Van Nguyen‡ Hendrik Schmitz§
DICE, University of DICE, University of University of Paderborn
Düsseldorf & CINCH Düsseldorf & CINCH CINCH & RWI

February 2015

Abstract

Since 2009, German nursing homes have been evaluated regularly with quality report cards published online. We argue that most of the information in the report cards does not reliably measure quality of care, but a subset of seven measures does. Using a sample of more than 3,000 nursing homes with information on two waves, we find a significant improvement in the nursing home quality from the first to the second evaluation. Both indicators comprising either the two outcome quality measures or the seven measures indicating “risk factors” in the report cards improve. This can be interpreted as evidence that quality disclosure positively affects the (reported) quality in nursing homes.

JEL Classification: L15, I11, I18

Keywords: public reporting, quality, long-term care, information

†We thank Florian Heiß, Ulrich Heimeshoff, and Christian Pfarr as well as participants of the DICE brown-bag seminar 2014, the CINCH Academy 2014 and the DIBOGS workshop in Fürth 2014 for valuable comments. Financial support from the BMBF (Förderkennzeichen 01EH1102A) is gratefully acknowledged.
‡Email: herr@dice.hhu.de; Address: Universität Düsseldorf, Düsseldorf Institute for Competition Economics (DICE), Universitätsstrasse 1, 40225 Düsseldorf, Germany. Tel: +49 211 81 15497, Fax: +49 211 81 15499
§Email: thu-van.nguyen@wiwinf.uni-due.de; Address: CINCH - Health Economics Research Center, Edmund-Körner-Platz 2, 45127 Essen, Germany. Tel: +49 201 183 2539
§University of Paderborn, Warburger Strasse 100, 33098 Paderborn, Germany, Tel.: +49 5251 603213, E-mail: hendrik.schmitz@uni-paderborn.de
1 Introduction

In this paper we analyze the effect of mandatory quality disclosure on outcome quality in the German nursing home market. After a series of public scandals regarding very bad quality in some German nursing homes (Tscharnke, 2009), health insurance providers and nursing home-owners took joint action to improve the transparency of nursing home quality. According to the “care transparency agreement (CTA),” started in 2009, German nursing homes are evaluated unannouncedly on a regular basis according to a standardized list of criteria. Since then, the quality information has been published online in standardized report cards and can be easily accessed on central websites.

Theoretically, the public provision of quality information should serve as an incentive for the quality improvement of nursing home providers. While, typically, there is an information asymmetry between providers and consumers on the true quality of care, public report cards reduce these asymmetries, giving consumers a higher bargaining and decision making power (Arrow, 1963). Good evaluations serve as a quality signal and consequently lead to a higher purchase probability and to a more efficient allocation (Reinstein and Snyder, 2005; Friberg and Grönqvist, 2012).

It has been empirically shown that public reporting increases quality in industries which deal with information asymmetries such as food labeling (Nielsen, 2006) or law schools (Stake, 2006). Regarding health care, some US and an Italian study find that public reporting leads to a quality improvement in hospitals (Laschober et al., 2007; Pham et al., 2006; Dziuban et al., 1994; Renzi et al., 2012). Looking at German hospitals, Busse et al. (2009) and Filistrucchi and Ozbugday (2012) identify an improvement in quality after public reporting. Cutler et al. (2004) argue that for evaluating the quality development of the hospital market, one potential problem could be the inaccuracy of data: better outcomes after quality information may also be due to the physicians’ selection of healthier patients. Thus, the question arises as to whether more information is always better (Dranove et al., 2003).

Turning to the nursing care market in the US, selected quality measures improved at least for subgroups of nursing homes following the introduction of the obligatory disclosure policy in 2002 (Lu, 2012; Park and Werner, 2011; Mukamel et al., 2008; Grabowski and Town, 2011). The policy was introduced by the Nursing Home Quality Initiative (NHQI)
in order to facilitate the search for an appropriate nursing home for consumers, as the US disclosure policy was not viewed as being sufficiently consumer-friendly (Stevenson, 2006; Kane and Kane, 2001). Werner et al. (2009b), who look at post-acute care, find that two out of three measures improve after public reporting. Along these lines, Werner et al. (2012) show that public reporting leads to improvements both in quality and in consumers choosing better-performing nursing homes in the US from 2001 to 2003. However, some sorting mechanism could also be a consequence of public reporting: Werner et al. (2011) show that high-risk patients tend to choose nursing homes that have achieved better results, whereas low-risk individuals rather choose low-scoring nursing homes.

Several channels might lead to quality improvements: First, that individuals choose better-performing nursing homes (“voting-by-feet”). The increased competition may lead to better quality (Werner et al., 2012). Grabowski and Town (2011), show using NHC that quality only increases in those nursing homes facing competitive pressure. However, a recent study on English nursing homes finds that competition reduces prices and quality (Forder and Allan, 2014). Competition is not the focus of this study.

Second, that nursing homes learn about their relative rank in terms of quality compared to others and may thus feel the need to improve. Third, better nursing homes may be able to attract more investors or to negotiate higher prices with the nursing care funds that may, again, increase quality. However, other mechanisms may be at work, too. In Germany, bad quality nursing homes are evaluated more frequently, possibly increasing the pressure to improve quality. Finally, reported quality might improve due to “teaching to the test” effects. Since nursing homes know the questions, they can prepare to score in exactly these aspects. Lu (2012) shows that unobserved quality measures may not improve, which gives rise to the “multitasking theory.” If resources are scarce and mainly put into the reported quality outcomes, less resources are left for the unobserved quality. Analogously, Werner et al. (2009a) provide some evidence for the improvement of reported quality, whereas they, too, state that the effect on unreported measures may be unclear. Lastly, in the US, the construction that residents only enter the quality measures in NHC after a minimum length of stay leads to a higher number of rehospitalizations of high-risk individuals before that threshold (Konetzka et al., 2013). Public reporting may thus lead to selection.
Quality increases also affect positively the financial performance of the improving nursing homes, independent of the former level (Park et al. (2011) using the 2002 NHC scores). Looking at prices, Clement et al. (2012) find price and quality increases after NHC for low-quality nursing homes only, while – using the new Five-Star-Rating – Hirth and Huang (2014) show that the well-rated nursing homes increase prices more than low-level nursing homes after public reporting. They conclude that consumers are responsive to public reporting, which can indeed enhance market efficiency.

In 2009, the Five-Star-Rating-System was introduced in the US and – similar to the German CTA – aims at higher transparency and facilitation of the decision making process of potential residents. Three aspects are included in the rating: health inspections, staffing, and quality measures. However, the rating system has been under debate recently, since only the health inspections are performed by the government, whereas the other two are reported by the nursing homes themselves. This may lead to an overvaluation, as the self-indicated reports are not audited by an external institution. In contrast to the US system, the evaluation process of the German disclosure policy is carried out by an independent institution. More detailed information can be found in Section 2.

Literature on nursing home quality in Germany is rather scarce. Some studies have examined the relationship between nursing home prices and quality (Mennicken, 2013; Reichert and Stroka, 2014) or have investigated the remuneration rates in nursing homes (Mennicken et al., 2014). Schmitz and Stroka (2014) analyze how consumers react to nursing home quality. The effect of higher transparency due to public reporting on nursing home quality in Germany has not been analyzed yet.

We use a sample of more than 3,000 German nursing homes which had been evaluated at least twice between 2009 and 2013. In analyzing the change in quality we focus on only two to seven of the 64 quality indicators in the report cards – the two that measure outcome quality and five more that are assumed to capture the “risk factors” shown below. The remaining ones mainly measure processes and services and are arguably uninformative about quality. We find that nursing homes indeed perform significantly better in the second wave than in the first.

This paper contributes to the literature in the following ways: it is the first study that measures the impact of higher transparency on the quality of German nursing homes and hereby exploits the panel structure of the German quality reports for the first time. In contrast to other studies, our constructed quality measures are mainly objective, do not depend on supply-side or demand-side characteristics, and are based on evaluations by an external institution (see discussion in Dranove et al., 2003 or Cutler et al., 2004).

This paper is organized as follows: Section 2 provides some information on the institutional background of the German nursing care market and the introduction of the report cards. Section 3 presents the data used. Section 4 analyzes changes in quality due to the report cards while robustness checks are reported in Section 5. Section 6 concludes.

2 Institutional background and quality assessments

The organization of long-term care is a self-administrative issue. Contracts between providers of the approximately 12,000 nursing homes and residents in need of care are individually agreed on. In principle, there is no regulated upper limit for the price that nursing homes may charge and nursing homes are to a large extent independent. However, prices cannot be set freely but result from a bargaining process between providers and sickness funds (see Schmitz and Stroka, 2014). Nursing homes are mostly run by non-profit (55%) or private (40%) institutions, while only 5% were public in 2009.

To ensure that all those in need can afford long-term care, the German long-term care insurance is obligatory and directly linked to the health insurance system, implying that almost everybody is covered by long-term care insurance (around 90 percent in the public and 10 percent in the private system). Formal care is partly financed by the health plan and partly out-of-pocket. For example, for the exemplary home in Figure A1 residents have to pay an additional €1,252 (care level I) or €1,845 (care level III) per month on top of the health plan’s coverage of €1,023 or €1,550, respectively, plus for any possible further special services or wishes (for a deeper analysis compare Section 5).
2.1 Quality assessment in German nursing homes

The Medical Review Board of the German Statutory Health Insurance (MRB) is responsible for monitoring quality in nursing homes, and as such serves as an external control body. Prior to 2008, the social long-term care system did not address the issue of quality reporting, as quality issues were dealt with bilaterally between the insurance and the service provider. The flaws of this approach were: that (1) unannounced quality evaluations by the MRB were not mandatory; and (2) the information was very difficult for the public to access.

Since 2008, the “care transparency agreement (CTA)” (Pflege-Transparenzvereinbarung)\(^2\) has been helping individuals in need of nursing home care to make a more informed nursing home choice. Comparability of nursing homes is guaranteed because the same 64 criteria are tested in all nursing homes and reporting of the results is standardized. The results of each evaluation are published not only in online report cards\(^3\) but are also displayed in the nursing homes.

Furthermore, as before, the implementation of the negotiated quality standards is obligatory: nursing homes are supposed to be punished should they not meet the requirements. However, the procedures are now more structured and follow specific guidelines. In this new setting, for instance, nursing homes may be forced to improve their failed standards until a repeated evaluation in the near future, where the fact that the nursing home had been evaluated again due to issues in the first evaluation will also be marked on the updated report card. We drop the two homes with replaced report cards from our data. The evaluation is unannounced and undertaken by trained inspectors of the respective regional MRB. All German nursing homes were finally tested at least once by the year 2011, followed by regular updates thereafter.

\(^2\)The public report cards were jointly set up by umbrella organizations of both health and long-term care insurances and owners of nursing homes (Bundesarbeitsgemeinschaft der überörtlichen Träger der Sozialhilfe, Bundesvereinigung der kommunalen Spitzenverbände, Vereinigung der Träger der Pflegeeinrichtungen and GKV-Spitzenverband (Central Organization of Care Insurance Funds, Associations of Care Institutions Sponsors at Federal Level, Federal Working Group of the Non-Local Sponsors of Social Benefits, and the Federal Association of Local Authority Central Organizations).

2.2 Criteria and average grades in the report cards

The criteria of the report cards comprise a wide range of aspects such as quality of care, handling of residents with dementia, quality of board and lodging, hygiene, as well as cultural offers. The full list of questions is reported in Table A1 in the Appendix. Consider, as an example, the most important criterion “Is the liquidity status of the resident appropriate?.” The inspectors test on a subgroup of residents in the nursing home, say 10 people, whether this criterion is fulfilled and calculate the percentage of individuals for whom it holds true. Then, the percentage value is translated into a grade according to the German system of school grades from 1.0 (= excellent) to 5.0 (= inadequate or failed) (see Table A2 in the Appendix for the mapping). The grades, not the exact percentage values, are then published. Many criteria are actually comprised by binary indicators (e.g. “Is there a systematic pain assessment?”’, where 1.0 stands for yes and 5.0 for no).

Since comparisons over 64 grades is rather difficult, an overall grade of the nursing home is generated by calculating an average of all single grades. Compare Figure A1 in the Appendix for a showcase report card which includes the overall grade and grades of four subgroups.

While there is no doubt that aggregation strongly facilitates the comparison, the aggregation method is subject to a great deal of critique among nursing scientists (see, e.g., Hasseler and Wolf-Ostermann, 2010). First, the mapping into school grades is arbitrary. It is highly disputable that fulfilling a criterion which is supposed to be standard for good quality only in, say, 75 per cent of all cases is a “good” quality (grade of 2.3, which represents a “good” in the German system). This obviously reflects the fact that the mapping is the result of an extensive bargaining process between the MRB and the nursing home owners before the care transparency agreement became effective. Second, averaging all 64 grades into the overall grade is problematic. More important criteria like outcome quality measures get the same weight as arguably less decisive factors like the offer of cultural activities in the nursing home. Apart from that, there is the critique that too much process and structural quality is measured but too little outcome quality (Hasseler and Wolf-Ostermann, 2010). Thus, we are sceptical of the content of the officially aggregated grades. That being said, we are nevertheless confident that the reports include a
lot of information on the quality of nursing homes, which we exploit to construct grades that reflect the true quality of care.

### 2.3 Extracting quality information from the report cards

In order to assess true quality improvement, we construct two quality indicators using single items from the reports. In defining our main quality indicator we use the only two outcome quality criteria among the 64 grades. These are the following (questions 15 and 18 in the report cards, see Table A1, own translations from German):

1. Is the nutritional status appropriate given the conditions set by the institution?
2. Is the supply of fluids appropriate given the conditions set by the institution?

In addition, we acknowledge that only a grade of 1.0 (excellent) implies that the criterion is fulfilled for all residents, truly reflecting a good quality of care. Therefore, we define binary indicators $q_j$ for criterion $j$ equaling one if and only if the criterion is fulfilled for all tested residents (grade 1.0) and zero if the grade is worse, meaning that it is not fulfilled for at least some residents:

$$q_k = \begin{cases} 
1 & \text{if grade}_j = 1.0 \\
0 & \text{if grade}_j > 1.0 
\end{cases}$$

We then define the quality indicator as

$$\text{Outcome quality} = \frac{1}{2} \sum_{j=1}^{2} q_j \quad j = 1, 2$$

Among all grades we consider *Nutritional and liquid status* as most important for the quality of care in a nursing home. Moreover, we argue that these are not easily manipulated in the short run for the mere purpose of an evaluation. We thereby address the potential issue that nursing homes might rather improve simple or cheap aspects which are, however, irrelevant for care quality, to increase the grade average.

In addition, we define a second indicator, measuring general quality of care. For this we follow the definition of Hasseler and Wolf-Ostermann (2010) and only use the seven “risk
criteria” instead of the full number of available grades to define an aggregate measure of quality of care in a nursing home:  

1. Is the nutritional status appropriate given the conditions set by the institution?
2. Is the supply of fluids appropriate given the conditions set by the institution?
3. Are documents regarding the treatment of chronic wounds or bedsores analyzed and, if necessary, the measures adjusted?
4. Are systematic pain assessments conducted?
5. Are individual risks and resources of residents with incontinence or a bladder catheter assessed?
6. Is the individual risk of contracture collected?
7. Do measures restricting the individual freedom require consent?

According to Hasseler and Wolf-Ostermann (2010), risk criteria are “factors that, when left unattended, affect the health and quality of life of individuals independent of the affliction.” (own translation of their originally German definition). We use a second indicator spanning all seven risk criteria, defined as

\[
\text{Care quality} = \frac{1}{7} \sum_{k=1}^{7} q_k \quad k = 1, \ldots, 7
\]

which is the share of ones among all seven criteria and called Care quality.  

While we do believe in the content of Care quality, it might be problematic as some of the indicators could be quickly improved without actually improving the true quality. Hence, Outcome quality is the preferred measure of quality.

---

4These are criteria 11, 15, 18, 20, 22, 27, 29 (order changed here). Note that 15 and 18 are the two outcome quality indicators also previously used. The care transparency agreement of December 17, 2008 scheduled a scientific evaluation of the report cards. An advisory board mainly composed of nursing and health scientists lead by Martina Hasseler and Karin Wolf-Ostermann was in charge of critically evaluating the report cards and making suggestions for improvement in future years. We follow some of their arguments published in the final report (Hasseler and Wolf-Ostermann, 2010).

5In order to avoid potential bias due to missing values, Care quality is redefined, as question numbers 11 and 29 have a high number of missing values: for the first wave, question 11 has about 1,255 missing values and question 29 has 674. For the second wave, there are 1,434 (question 11) and 801 (question 29) missing values. Therefore, if one or both of them is a missing value, the share accordingly reduces to \((\frac{0}{7}, \frac{1}{7}, \ldots, \frac{2}{7}, \frac{3}{7})\) and \((\frac{0}{7}, \frac{1}{7}, \ldots, \frac{2}{7})\), respectively. These outcomes are then mapped on the original scale to the closest neighboring value.
3 Data and descriptive statistics

3.1 Report cards

The unit of observation is the nursing home. We merge three data sets: the report cards of German nursing homes, additional pricing information for 2009, and regional information at county level. The report cards are available online for all nursing homes in Germany, which amount to about 12,000. However, for the first wave, we only have access to a random sample of roughly 5,000 nursing homes and no price information. We thus add the latter from the nursing database PAULA (Pflege Angebot und Leistungsanbieter (2010), data of the BKK Bundesverband). We exclude nursing homes only providing short-term and out-patient care and care for children and disabled individuals. We also exclude nursing homes that are specialized in only treating residents suffering from dementia. Various special homes such as care for residents with apallic conditions, multiple sclerosis, or stroke residents are also dropped. Finally, we exclude nursing homes with less than 10 residents. Thus, we only focus on general long-term care. As we want to exploit the panel structure, we only include homes that were observed in both waves. In total, we end up with 6,176 nursing home-year observations, 3,088 observations for each wave. The nursing homes have been evaluated at different points in time: 2,769 in 2009 and 319 in 2010 in the first wave, 135 in 2011, 2,670 in 2012, and 283 in 2013 in the second wave (see Figure 1).

Figure 1: Number of observations at quarter of evaluation
Table 1 reports the descriptive statistics of the two quality measures in the first and the second wave separately. In both cases it can be seen that nursing homes have improved over time.

Table 1: Descriptives: Constructed quality measures in 1st and 2nd evaluation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcome Quality</td>
<td>0.799</td>
<td>0.328</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Care Quality</td>
<td>0.582</td>
<td>0.257</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2nd Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcome Quality</td>
<td>0.912</td>
<td>0.216</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Care Quality</td>
<td>0.739</td>
<td>0.219</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Number of observations: 3,088 each wave. Sources: Care quality and outcome quality: own calculations using data from the transparency reports.

In the first evaluation, only 80 percent of the nursing homes manage to fulfil both, the provision with sufficient nutrition and with liquids, of all selected inhabitants. Care quality can also be considered as moderate in the first evaluation, as roughly 58 percent of the seven relevant criteria are fulfilled, and only 347 out of a total of 3,088 nursing homes fulfil all care quality criteria (not shown in the table). In the second wave, the nursing homes achieve better results across all criteria: outcome quality increases to 91 percent and the number of nursing homes which fulfil all seven care quality criteria more than doubles to 750. On average, nursing homes now fulfil 74 percent of the care criteria, which is equivalent to meeting five of the seven requirements.

3.2 Regional characteristics at county level

Socio-economic control variables measured at the county level are taken from the Federal Office for Building and Regional Planning (INKAR) for the years 2009 to 2011 (only these two years are available) and we match the first wave to 2009 and the second to 2011. Descriptive statistics in Table 2 show that in our sample, on average, net household income amounts to roughly €1,530 per month. Furthermore, population and nursing-care specific factors are included, such as the share of elderly above 65 years old, the ratio of care benefit receivers to people in need of care, and the amount of the average pension payment per month. The average pension payment per month of men (€1,073) is almost twice as high as the women’s pensions (€566). Finally, the percentage of a particular
county defined as rural as well as federal state dummies are also included in our main regressions.

Table 2: Descriptives: Regional characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household income [in 1,000 EUR]</td>
<td>1.51</td>
<td>0.199</td>
<td>1.082</td>
<td>2.45</td>
</tr>
<tr>
<td>Share of population 65+</td>
<td>0.209</td>
<td>0.023</td>
<td>0.15</td>
<td>0.285</td>
</tr>
<tr>
<td>Ppl in need of care [per 10,000 inhab.]</td>
<td>302.189</td>
<td>61.324</td>
<td>154.8</td>
<td>542.6</td>
</tr>
<tr>
<td>Recipients LTC allowance per ppl in need</td>
<td>0.457</td>
<td>0.064</td>
<td>0.275</td>
<td>0.692</td>
</tr>
<tr>
<td>Pension Payment (m) [in 1,000 EUR]</td>
<td>1.047</td>
<td>0.08</td>
<td>0.829</td>
<td>1.332</td>
</tr>
<tr>
<td>Pension Payment (f) [in 1,000 EUR]</td>
<td>0.558</td>
<td>0.1</td>
<td>0.37</td>
<td>0.796</td>
</tr>
<tr>
<td>Share of county defined as rural</td>
<td>0.205</td>
<td>0.251</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Physicians per 10,000 inhab.</td>
<td>16.658</td>
<td>5.2</td>
<td>8.239</td>
<td>39.185</td>
</tr>
<tr>
<td>Land price [EUR/m²]</td>
<td>156.201</td>
<td>143.563</td>
<td>4.748</td>
<td>1076.682</td>
</tr>
</tbody>
</table>


4 Quality responses to the transparency reform

4.1 Estimation strategy

We postulate that the transparency reform, especially public reporting, has a positive impact on the suppliers’ behavior, incentivizing them to put more effort into improving quality. Thus, in our empirical analysis, we are mostly interested in the change in quality measures between the first and the second wave of published report cards. As the transparency reform is applied to all German nursing homes, the analysis is essentially a before-and-after comparison. We assume that the first evaluation measures the baseline quality that would have also been prevalent – but not measured and published – without the report cards. Reactions to the publications should be visible in the second evaluation. We estimate the following linear model:

\[
\text{quality}_{it} = \beta_0 + \beta_1 \text{2nd evaluation}_{it} + X_{it}\delta + \tau_t + \lambda_{FS_i} + \epsilon_{it}
\]  

where \(\text{quality}_{it}\) is either \(\text{Outcome quality}\) or \(\text{Care quality}\). The main explanatory variable of interest is \(\text{2nd evaluation}\) which is a dummy variable equalling one in the second wave
and zero in the first. As we also include year fixed effects $\tau_t$, $\beta_1$ measures the discontinuous jump in quality between both waves. The year fixed effects are supposed to mainly capture the learning effects of the evaluators over time. The results below do not change if a linear trend or no time trend at all are included instead. $X_{it}$ is a vector containing the information on county level shown in Table 2 and $\lambda_{FS_i}$ contains 15 federal state fixed effects. Standard errors are clustered on nursing home level.

To identify the effects of the reform on quality, we need two main assumptions.

1. Without public reporting, the nursing homes would not have changed their quality.

2. The indicator $quality_{it}$ indeed measures the quality of nursing homes. That is, changes in the outcome variable reflect true quality changes.

### 4.2 Results

Table 3 reports the regression results. At the second evaluation, nursing homes have, on average, improved their outcome quality (sufficient provision with liquids and nutrition) by 0.125. This is equivalent to the case where 12.5 percent of all nursing homes did not satisfy the sufficient nutritional and liquid requirements during the first evaluation and changed this in the meantime. Regarding care quality, the transparency reform led to an improvement by 13.5 percentage points, which is equal to roughly one additionally fulfilled criterion out of seven per nursing home (since $1/7 = 0.143$). Considering the fact that about 350 nursing homes had fulfilled all criteria in the first wave, this number even more than doubled at the second evaluation to 750. Nursing homes seem to have improved in several aspects, as the number of those fulfilling none of the seven criteria decreased from 74 to merely six in the course of the transparency reform, whereas the number of nursing homes fulfilling six out of seven criteria increased from 452 (1st evaluation) to 856 (2nd evaluation).

The coefficient of the indicator for year 2010 ($Year_{2010}$) is significantly positive, which means that within the first wave nursing homes which were evaluated first perform worse than nursing homes which were evaluated later. This pattern does not occur for the second wave.
<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome Quality</strong></td>
<td><strong>Care quality</strong></td>
</tr>
<tr>
<td>2nd evaluation</td>
<td>0.123***</td>
</tr>
<tr>
<td>(0.016)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Evaluated in 2010</td>
<td>0.066***</td>
</tr>
<tr>
<td>(0.017)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Evaluated in 2012</td>
<td>-0.017</td>
</tr>
<tr>
<td>(0.015)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Evaluated in 2013</td>
<td>0.012</td>
</tr>
<tr>
<td>(0.019)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>Household income [in 1,000 EUR]</td>
<td>0.002</td>
</tr>
<tr>
<td>(0.030)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Share of population 65+</td>
<td>0.342</td>
</tr>
<tr>
<td>(0.224)</td>
<td>(0.195)</td>
</tr>
<tr>
<td>Recipients LTC allowance per ppl in need</td>
<td>0.066</td>
</tr>
<tr>
<td>(0.089)</td>
<td>(0.076)</td>
</tr>
<tr>
<td>Pension Payment (m) [in 1,000 EUR]</td>
<td>-0.123</td>
</tr>
<tr>
<td>(0.076)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>Pension Payment (f) [in 1,000 EUR]</td>
<td>-0.082</td>
</tr>
<tr>
<td>(0.105)</td>
<td>(0.087)</td>
</tr>
<tr>
<td>Share of county defined as rural</td>
<td>-0.004</td>
</tr>
<tr>
<td>(0.024)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.895***</td>
</tr>
<tr>
<td>(0.102)</td>
<td>(0.083)</td>
</tr>
<tr>
<td>Federal States FE</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>6176</td>
</tr>
<tr>
<td>R²</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Standard errors clustered on nursing home level.
The regional characteristics do not explain much of the variation in nursing home quality, especially as federal state fixed effects are included. A higher quality can be found in counties with a higher share of elderly and of benefit receivers and with lower pension payment for men given the household income and only looking at care quality.

Summing up, both grades show a significant improvement between both waves. Hence, nursing homes react to the once-experienced unannounced evaluation by offering better quality. Certainly, it is not clear whether the effect stems from some kind of “teaching to the test” phenomenon, such that nursing homes only improve in exactly those criteria which they know will be checked (Lu, 2012). However, we argue that even if this were the case, this would still be an improvement in very important quality criteria (seven risk factors and in particular the nutritional and liquid status of residents) and, therefore, beneficial to residents. Public reporting, therefore, may serve as an instrument in order to steer care providers toward investing in better quality, as they do react to the public provision of quality information.

5 Robustness checks

5.1 Accounting for the size and the price of a nursing home

Potentially, the quality of a nursing home also depends on its size and its price level. However, these two variables are possibly endogenous due to reversed causality (prices and demand, thus, size in the long run, depend on the nursing home’s quality). In principle it is, therefore, preferred to leave them out of the regression model as long as both potentially endogenous variables are uncorrelated with the main explanatory variable.

In a robustness check, we include both price and size in the linear regression and estimate IV models where both are instrumented separately. Note, however, that the indicator for the second evaluation is not instrumented as this is unlikely to be endogenous. The main aim of this analysis is to test whether the results from the previous section are robust to including these two variables.

Table A3 in the Appendix shows descriptive statistics of the endogenous variables. The number of residents is around 80 on average (ranging from 10 to 812). Depending on
the care level, the social LTC insurance covers only part of the total care costs. Thus, we include the personal contribution of the price individuals have to pay out-of-pocket, since this is the relevant price from the demanders’ point of view. The long-term care insurance covers, for example in North-Rhine Westphalia in 2013, between €1,023 (care level I) and €1,550 (care level III) per month. Individuals have to pay additionally roughly €1,095 on average out-of-pocket, with a minimum of €142 and a maximum amount of €2,285 per month. Although prices depend on the resident’s care level, we use simple mean prices per home instead, as we do not have information on the distribution of the various care levels in each nursing home.

Here, we drop all 319 nursing homes evaluated in 2010 and 783 in total due to missing price information. Prices are deflated, using 2009 as the base year.

To account for possible reverse causality, we instrument the number of residents and the out-of-pocket price separately by two instruments: the number of physicians per 10,000 inhabitants and the average land price in the respective county. Physician density is strongly correlated with the number of residents and the nursing home price, since physicians are more likely to settle in higher populated areas and it is more attractive and more expensive to live in urban than in rural areas. However, there should not be a direct effect of the number of physicians on nursing home quality. The same applies to land price, which is highly correlated with the nursing home prices and the number of residents, as more popular areas are generally more expensive and more attractive, but land price should not have a direct effect on nursing home quality.

Table 4 summarizes the results for both quality measures. The first column (1) repeats the main results from Table 3. Columns (2) to (4) show the extended models, (2) including the endogenous variables without instrumenting them, (3) instrumenting the number of residents and excluding price information, and finally (4) instrumenting price, excluding the number of residents. The same estimations are performed for care quality. As can be seen, the positive effects of the second evaluation do not differ much in terms of magnitude and significance, and thereby remain robust throughout all estimation strategies. The full estimation results in Table A4 report that size and prices are not significant once they are instrumented.
Table 4: Overview of main results (OLS) versus three robustness checks

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS w/o size and price</td>
<td>OLS including size and price</td>
<td>IV including size</td>
<td>IV including size and price</td>
</tr>
<tr>
<td>Outcome quality</td>
<td>2nd evaluation</td>
<td>0.123***</td>
<td>0.108***</td>
<td>0.111***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.018)</td>
</tr>
<tr>
<td></td>
<td>F-test 1st stage</td>
<td>-</td>
<td>-</td>
<td>18.52</td>
</tr>
<tr>
<td>Care quality</td>
<td>2nd evaluation</td>
<td>0.133***</td>
<td>0.122***</td>
<td>0.125***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.020)</td>
</tr>
<tr>
<td></td>
<td>F-test 1st stage</td>
<td>-</td>
<td>-</td>
<td>18.52</td>
</tr>
<tr>
<td>County level controls</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Year and federal state FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

Number of observations in each regression: 6,176 (3,088 each wave). Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Instruments: land price and physician density per 10,000 inhabitants. IV regressions: using the ivreg2 command for Stata (Baum et al., 2002). Full estimation results reported in Table A4 in the Appendix.

5.2 Potential sample selection

We only include nursing homes that appear twice in our data set. In order to test for sample selection, we analyze whether our sample differs significantly in quality from those nursing homes we observe only once. When we apply a probit model to estimate the probability of a nursing home being observed more than once including the quality measures, the year of evaluation, federal state dummies as well as regional characteristics as explanatory variables we can show that quality is not associated with being selected into our sample. None of the two quality measures’ coefficients is significantly different from zero (see Table A5 in the Appendix). Thus, there seems to be no systematic entering into or dropping out of the sample.

6 Conclusion

This study analyzes the effects of a reform to increase transparency in health care: the mandatory evaluation by an external institution and publication of these report cards for
German nursing homes. We use a random sample of more than 3,000 German nursing homes evaluated twice between 2009 and 2013. Our results show that the higher transparency indeed has a positive impact on the reported quality of nursing homes. The evaluated nursing homes, on average, increased their performance by one out of seven grade units from the first to the second evaluation.

This can be seen as a positive effect of the transparency reform. However, it remains unclear whether this is just a “teaching to the test” effect or not. We argue that, even if this were the case, it is an indication of the reform’s success as improvements in the two outcome quality indicators are certainly beneficial for the residents. This holds as long as nursing homes do not shift resources away from other important but not tested outcomes.

Similar to the US, the evaluation process is subject to current debates, as the average grades could be misleading due to a potential overvaluation. However, the reasons for the debates in the two countries differ. While the Five Star Rating suffers from self-reporting, the German averages may hide specific quality issues. The most important single indicators are reported but hard to identify in the report cards and individuals looking for suitable nursing homes may focus on the easy to interpret overall grade. Moreover, the chosen increments of each grade, which reflect the result of a bargaining process between insurance companies and nursing homes, are too good, on average (see the discussion in Section 2). Hence, several adjustments still need to be made. We argue that the vast majority of indicators cannot be used to measure quality – only seven indicators can.

Note, however, that we do not make any statement about the current level of the quality in German nursing homes, only on changes.

Therefore, the report cards should be strongly revised to include many more outcome quality indicators and indicators of quality of life of the residents. Moreover, the mapping from indicators to grades should be more realistic, meaning that poor quality should not be labelled “good” or even “very good.” The results of this study allow us to infer that such a reform would most likely lead to quality improvements in the German nursing home sector.
References


Appendix

Figure A1: Example of a report card (first page)

Quality of nursing home
NOVA VITA Residenz Essen

Goethestr. 19, 45148 Essen · Tel.: 0201 / 2901 · Fax: 0201 / 292700
mkv@novavita.com · www.novavita.com

Overall result
Patient questionnaire
State average
2.2 good
2.7 satisfactory
1.3 excellent
1.4 excellent
2.0 good
1.7 good
1.0 excellent

MDK-Evaluation

<table>
<thead>
<tr>
<th>Erklärun gen zum Bewertungssystem</th>
<th>Kommentar der Pflegeeinrichtung</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDK-Prüfung am: 08.10.2009</td>
<td>weitere Leistungsgarantien und Strukturdaten</td>
</tr>
<tr>
<td>Anzahl der versorgten Bewohner: 53</td>
<td></td>
</tr>
<tr>
<td>Anzahl der in der Prüfung einbezogenen Bewohner: 6</td>
<td></td>
</tr>
<tr>
<td>Anzahl der behandelten Personen: 5</td>
<td></td>
</tr>
<tr>
<td>Die Pflegeeinrichtung hat eine</td>
<td></td>
</tr>
<tr>
<td>Wiederholungsprüfung beantragt:</td>
<td></td>
</tr>
</tbody>
</table>
| Pflegeeinrichtung im Bundesland: | |}

Source: www.pflegenoten.de. Own translations to English. Only the first page is shown. The remaining pages include all 64 single criteria.
### Area 1: Nursing and medical care

1. Is an active communication with a physician comprehensible if required?
2. Does the application of the nursing treatments correspond to the physician’s orders?
3. Does the supply of medicines correspond to the physician’s orders?
4. Is the use of medicines appropriate?
5. Are compression stockings put on properly?
6. Is the individual pressure sore risk being assessed?
7. Are pressure ulcer prevention measures being applied?
8. Are place and time at which the chronic wound/pressure ulcer occurred verifiable?
9. Is a differentiated documentation in case of chronic wounds or pressure ulcer being carried out (in terms of actuality, verifiability of development, size, position, depth)?
10. Are the applied measures to treat chronic wounds or pressure ulcer based on state-of-the-art knowledge?
11. Are documents regarding the treatment of chronic wounds or bedsores analyzed and, if necessary, the measures adjusted?
12. Do residents with chronic pains receive the prescribed medication?
13. Are individual nutritional resources and risks documented?
14. Are necessary measures taken in case of restrictions regarding independent supply of food?
15. Is the nutritional status appropriate given the conditions set by the institution?
16. Are individual resources and risks regarding the supply of fluids documented?
17. Are necessary measures taken in case of restrictions regarding independent supply of fluids?
18. Is the supply of fluids appropriate given the conditions set by the institution?
19. Is the sense of taste of residents with feeding tubes being stimulated?
20. Are systematic pain assessments conducted?
21. Does the nursing home cooperate closely with the treating physician?
22. Are individual risks and resources of residents with incontinence or a bladder catheter assessed?
23. Are necessary measures for residents with incontinence or a bladder catheter taken?
24. Is the individual risk of falling assessed?
25. Are fall incidents being documented?
26. Are necessary prophylaxes against fall incidents taken?
27. Is the individual risk of contracture collected?
28. Are necessary contracture prophylaxes taken?
29. Do measures restricting the individual freedom require consent?
30. Is the necessity of freedom restricting measures checked regularly?
31. Are individual needs and habits of the residents regarding personal hygiene taken into account and being carried out accordingly?
32. Are individual needs and habits of the residents regarding oral and dental hygiene taken into account and being carried out accordingly?
33. Is nursing care usually being carried out by the same nurse?
34. Are workers regularly trained regarding First Aid and emergency measures?
35. Do written procedural instructions regarding First Aid and emergency measures exist?

### Area 2: Care of residents suffering dementia

36. Is the biography of residents suffering dementia taken into account and being considered when planning daily activities?
37. Are accompanying and caring persons of residents suffering dementia incorporated into the nursing and caring process?
38. Is self-determinantion of residents suffering dementia taken into account in the nursing and caring process?
39. Is well-being of residents suffering dementia determined and documented, and appropriate measures for improvement deducted from that information?
Table A1 – Continued

[40] Do suitable exercise and recreational areas for particular target groups exist (at night time also)?

[41] Do secured recreational areas outside exist?

[42] Do identification facilitating arrangements regarding design of surroundings exist in rooms and recreation rooms?

[43] Are individual guidance measures, e.g., photographs, used?

[44] Are residents suffering dementia offered adequate activities, e.g., regarding exercise, communication, or perception?

[45] Are residents suffering dementia offered suitable food?

Area 3: Social care and the arrangement of the daily routine

[46] As part of social care, is group counseling available?

[47] As part of social care, is individual counseling available?

[48] Does the nursing home have annual celebrations?

[49] Are there activities together with the local community?

[50] Are there measures to promote contact with relatives?

[51] Are the social care measures justified by the residents’ composition and needs?

[52] Is assistance or information provided to familiarize new residents with the nursing facility (e.g., contact person, support during the orientation, assessment interviews after six weeks)?

[53] Is the orientation phase systematically evaluated?

[54] Are there guidelines with respect to the provision of terminal care?

[55] Does the nursing facility have a system for managing complaints?

Area 4: Accommodation, provision, household management, and hygiene

[56] Are residents allowed to decorate and design their rooms with their own furniture, personal effects, and memorabilia?

[57] Do residents have a say in the design and decoration of the communal areas?

[58] Does the facility give a good overall impression in terms of cleanliness and hygiene? For example, does it appear clean? Is it in order? Are there unpleasant odors?

[59] Within a specified time slot, are residents free to choose when to eat?

[60] Is appropriate food provided for people with special dietary requirements (e.g., residents with diabetes)?

[61] Is the food plan made available to the residents in a legible format?

[62] Is the presentation of food and drinks tailored to the needs of each individual resident? For example, to facilitate eating and digestion, some residents require food to be precut into smaller pieces or pureed.

[63] Are the portions tailored to the preferences of the residents?

[64] Are the food and drinks for the residents provided in a pleasant environment and relaxing atmosphere?

Note: Outcome and risk criteria highlighted.
### Table A2: Mapping of grades

| Category          | Grade | Percentage range | | Category | Grade | Percentage range |
|-------------------|-------|------------------||----------|-------|------------------|
| excellent quality | 1.0   | 97.4 - 100.0     | | 3.5      | 57.6 - 58.9      |
| (sehr gut)        | 1.1   | 94.8 - 97.3      | | 3.6      | 56.2 - 57.5      |
| good quality      | 1.2   | 92.2 - 94.7      | | 3.7      | 54.8 - 56.1      |
|                   | 1.3   | 89.6 - 92.1      | | poor     | 3.8      | 53.4 - 54.7      |
|                   | 1.4   | 87.0 - 89.5      | | quality  | 3.9      | 52.0 - 53.3      |
|                   | 1.5   | 85.6 - 86.9      | | 4.0      | 50.6 - 51.9      |
|                   | 1.6   | 84.2 - 85.5      | | 4.1      | 49.2 - 50.5      |
|                   | 1.7   | 82.8 - 84.1      | | 4.2      | 47.8 - 49.1      |
|                   | 1.8   | 81.4 - 82.7      | | 4.3      | 46.4 - 47.7      |
|                   | 1.9   | 80.0 - 81.3      | | 4.4      | 45.0 - 46.3      |
|                   | 2.0   | 78.6 - 79.9      | | 4.5      | 43.6 - 44.9      |
|                   | 2.1   | 77.2 - 78.5      | | 4.6      | 42.2 - 43.5      |
|                   | 2.2   | 75.8 - 77.1      | | 4.7      | 40.8 - 42.1      |
|                   | 2.3   | 74.4 - 75.7      | | 4.8      | 39.4 - 40.7      |
|                   | 2.4   | 73.0 - 74.3      | | 4.9      | 38.0 - 39.3      |
|                   | 2.5   | 71.6 - 72.9      | | 5.0      | 0.0 - 37.9       |
|                   | 2.6   | 70.2 - 71.5      | | 2.7      | 68.8 - 70.1      |
|                   | 2.8   | 67.4 - 68.7      | | 2.9      | 66.0 - 67.3      |
|                   | 3.0   | 64.6 - 65.9      | | 3.1      | 63.2 - 64.5      |
|                   | 3.2   | 61.8 - 63.1      | | 3.3      | 60.4 - 61.7      |
|                   | 3.4   | 59.0 - 60.3      | | 4.0      | 50.6 - 51.9      |
|                   | 4.1   | 49.2 - 50.5      | | 4.2      | 47.8 - 49.1      |
|                   | 4.3   | 46.4 - 47.7      | | 4.4      | 45.0 - 46.3      |
|                   | 4.5   | 43.6 - 44.9      | | 4.6      | 42.2 - 43.5      |
|                   | 4.7   | 40.8 - 42.1      | | 4.8      | 39.4 - 40.7      |
|                   | 4.9   | 38.0 - 39.3      | | 5.0      | 0.0 - 37.9       |
|                   | 5.0   | 0.0 - 37.9       | | 2.5      | 68.8 - 70.1      |


### Table A3: Descriptives: endogenous explanatory variables:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal Contribution</td>
<td>1081.728</td>
<td>311.628</td>
<td>150.14</td>
<td>3574.375</td>
</tr>
<tr>
<td>Overall NH Fees</td>
<td>2339.061</td>
<td>311.628</td>
<td>1407.473</td>
<td>4831.708</td>
</tr>
<tr>
<td># Residents</td>
<td>81.901</td>
<td>46.606</td>
<td>10</td>
<td>811</td>
</tr>
<tr>
<td>N</td>
<td>2,505</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal Contribution</td>
<td>1114.153</td>
<td>344.075</td>
<td>141.761</td>
<td>4214.230</td>
</tr>
<tr>
<td>Overall NH Fees</td>
<td>2330.402</td>
<td>344.446</td>
<td>1349.164</td>
<td>5431.292</td>
</tr>
<tr>
<td># Residents</td>
<td>83.337</td>
<td>46.559</td>
<td>10</td>
<td>812</td>
</tr>
<tr>
<td>N</td>
<td>2,888</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Transparency reports. Information given on nursing home level. Prices are deflated. Base year = 2009. Overall nursing home fees reported for completeness but not used in the regression analysis.
Table A4: Effect on outcome quality and care quality: Possibly endogenous factors instrumented

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd evaluation</td>
<td>0.108***</td>
<td>0.111***</td>
<td>0.116***</td>
<td>0.122***</td>
<td>0.125***</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.018)</td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Log(Personal Contribution)</td>
<td>0.057***</td>
<td>-0.118</td>
<td>0.077***</td>
<td>0.017</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.171)</td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Log(# residents)</td>
<td>-0.054***</td>
<td>-0.046</td>
<td>0.017</td>
<td>0.025</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.068)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Evaluated in 2012</td>
<td>-0.003</td>
<td>-0.004</td>
<td>-0.008</td>
<td>0.017</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.018)</td>
<td>(0.017)</td>
<td>(0.020)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Evaluated in 2013</td>
<td>0.024</td>
<td>0.024</td>
<td>0.023</td>
<td>0.054**</td>
<td>0.053**</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.020)</td>
<td>(0.021)</td>
<td>(0.024)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Household income [in 1,000 EUR]</td>
<td>-0.007</td>
<td>0.001</td>
<td>0.023</td>
<td>-0.016</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.035)</td>
<td>(0.037)</td>
<td>(0.026)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Share of population 65+</td>
<td>0.570***</td>
<td>0.486***</td>
<td>0.298</td>
<td>0.682***</td>
<td>0.528**</td>
</tr>
<tr>
<td></td>
<td>(0.227)</td>
<td>(0.228)</td>
<td>(0.330)</td>
<td>(0.198)</td>
<td>(0.209)</td>
</tr>
<tr>
<td>Recipients LTC allowance per ppl in need</td>
<td>0.078</td>
<td>0.098</td>
<td>0.131</td>
<td>0.291***</td>
<td>0.298***</td>
</tr>
<tr>
<td></td>
<td>(0.094)</td>
<td>(0.095)</td>
<td>(0.117)</td>
<td>(0.079)</td>
<td>(0.083)</td>
</tr>
<tr>
<td>Pension Payment (m) [in 1,000 EUR]</td>
<td>-0.113</td>
<td>-0.115</td>
<td>-0.127</td>
<td>-0.200***</td>
<td>-0.227***</td>
</tr>
<tr>
<td></td>
<td>(0.080)</td>
<td>(0.083)</td>
<td>(0.081)</td>
<td>(0.066)</td>
<td>(0.072)</td>
</tr>
<tr>
<td>Pension Payment (f) [in 1,000 EUR]</td>
<td>0.035</td>
<td>0.061</td>
<td>0.068</td>
<td>0.104</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(0.106)</td>
<td>(0.148)</td>
<td>(0.157)</td>
<td>(0.087)</td>
<td>(0.126)</td>
</tr>
<tr>
<td>Share of county defined as rural</td>
<td>-0.006</td>
<td>-0.007</td>
<td>-0.007</td>
<td>-0.054**</td>
<td>-0.047**</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.026)</td>
<td>(0.026)</td>
<td>(0.021)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.622***</td>
<td>0.981***</td>
<td>1.630</td>
<td>0.194</td>
<td>0.444**</td>
</tr>
<tr>
<td></td>
<td>(0.173)</td>
<td>(0.250)</td>
<td>(1.170)</td>
<td>(0.143)</td>
<td>(0.210)</td>
</tr>
<tr>
<td>Federal States FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>5391</td>
<td>5391</td>
<td>5391</td>
<td>5391</td>
<td>5391</td>
</tr>
<tr>
<td>R²</td>
<td>0.11</td>
<td>0.11</td>
<td>0.09</td>
<td>0.27</td>
<td>0.24</td>
</tr>
<tr>
<td>F first stage</td>
<td>18.52</td>
<td>29.74</td>
<td>15.82</td>
<td>18.54</td>
<td>29.74</td>
</tr>
</tbody>
</table>

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered on nursing home level. Outcome quality: (1) OLS with both endog vars, (2) IV for residents, (3) IV for price. Care quality: (4) OLS with both endog vars, (5) IV for residents, (6) IV for price. Instruments: land price and physician density per 10,000 inhabitants. IV regressions: using the ivreg2 command for Stata (Baum et al., 2002). Evaluated in 2010 dropped due to missing price information.
Table A5: Robustness Check: probability to be included in the balanced panel

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Care Quality</td>
<td>0.025</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>Outcome Quality</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.223</td>
<td>0.192</td>
</tr>
<tr>
<td></td>
<td>(0.330)</td>
<td>(0.331)</td>
</tr>
<tr>
<td>Federal State Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>County level controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>PR²</td>
<td>0.101</td>
<td>0.101</td>
</tr>
<tr>
<td>N</td>
<td>12,503</td>
<td>12,503</td>
</tr>
</tbody>
</table>

Significance levels: * p< 0.1 ** p< 0.05 *** p< 0.01. Standard errors clustered on nursing home level. Dependent variable = probability of being evaluated twice. We report coefficients of probit regressions. The sample includes nursing homes in both waves and is an unbalanced panel of all nursing homes with quality information.
PREVIOUS DISCUSSION PAPERS


174 Buchwald, Achim, Competition, Outside Directors and Executive Turnover: Implications for Corporate Governance in the EU, February 2015.

173 Buchwald, Achim and Thorwarth, Susanne, Outside Directors on the Board, Competition and Innovation, February 2015.


171 Haucap, Justus, Heimeshoff, Ulrich and Siekmann, Manuel, Price Dispersion and Station Heterogeneity on German Retail Gasoline Markets, January 2015.

170 Schweinberger, Albert G. and Suedekum, Jens, De-Industrialisation and Entrepreneurship under Monopolistic Competition, January 2015.

169 Nowak, Verena, Organizational Decisions in Multistage Production Processes, December 2014.


164 Caprice, Stéphane, von Schlippenbach, Vanessa and Wey, Christian, Supplier Fixed Costs and Retail Market Monopolization, October 2014.

163 Klein, Gordon J. and Wendel, Julia, The Impact of Local Loop and Retail Unbundling Revisited, October 2014.


160 Behrens, Kristian, Mion, Giordano, Murata, Yasusada and Suedekum, Jens, Spatial Frictions, September 2014.

Stiebale, Joel, Cross-Border M&As and Innovative Activity of Acquiring and Target Firms, August 2014.


Baumann, Florian and Friehe, Tim, On Discovery, Restricting Lawyers, and the Settlement Rate, August 2014.


Kholodilin, Konstantin A., Thomas, Tobias and Ulbricht, Dirk, Do Media Data Help to Predict German Industrial Production?, July 2014.


Jeitschko, Thomas D., Jung, Yeonjei and Kim, Jaesoo, Bundling and Joint Marketing by Rival Firms, May 2014.

Dauth, Wolfgang and Suedekum, Jens, Globalization and Local Profiles of Economic Growth and Industrial Change, April 2014.

Nowak, Verena, Schwarz, Christian and Suedekum, Jens, Asymmetric Spiders: Supplier Heterogeneity and the Organization of Firms, April 2014.

Hasnas, Irina, A Note on Consumer Flexibility, Data Quality and Collusion, April 2014.

Baye, Irina and Hasnas, Irina, Consumer Flexibility, Data Quality and Location Choice, April 2014.


117 Sapi, Geza and Suleymanova, Irina, Consumer Flexibility, Data Quality and Targeted Pricing, November 2013.


115 Baumann, Florian, Denter, Philipp and Frieh Tim, Hide or Show? Endogenous Observability of Private Precautions Against Crime When Property Value is Private Information, November 2013.


Baumann, Florian and Friehe, Tim, Competitive Pressure and Corporate Crime, September 2013.


Baumann, Florian and Friehe, Tim, Design Standards and Technology Adoption: Welfare Effects of Increasing Environmental Fines when the Number of Firms is Endogenous, September 2013.


93 Baumann, Florian and Friehe, Tim, Status Concerns as a Motive for Crime?, April 2013.


91 Baumann, Florian and Friehe, Tim, Private Protection Against Crime when Property Value is Private Information, April 2013. Published in: International Review of Law and Economics, 35 (2013), pp. 73-79.


88 Jovanovic, Dragan, Mergers, Managerial Incentives, and Efficiencies, April 2014 (First Version April 2013).


84 Bataille, Marc and Steinmetz, Alexander, Intermodal Competition on Some Routes in Transportation Networks: The Case of Inter Urban Buses and Railways, January 2013.


Older discussion papers can be found online at: http://ideas.repec.org/s/zbw/dicedp.html