DOI: 10.1111/j.1475-6773.2007.00803.x
RESEARCH ARTICLE

Nursing Homes

The Staffing–Outcomes Relationship in Nursing Homes

R. Tamara Konetzka, Sally C. Stearns, and Jeongyoung Park

Objective. To assess longitudinally whether a change in registered nurse (RN) staffing and skill mix leads to a change in nursing home resident outcomes while controlling for the potential endogeneity of staffing.
Data Sources. Minimum Data Set (MDS) nursing home resident assessment data from five states merged with Online Survey Certification and Reporting (OSCAR) data from 1996 through 2000.
Study Design. Resident-level longitudinal analysis with facility fixed effects and instrumental variables. Outcomes studied are incidence of pressure sores and urinary tract infections. RN staffing was measured as the care hours per resident-day and skill mix was measured as RN staffing hours as a proportion of total staffing hours.
Data Extraction Method. We use all quarterly MDS assessments that fall within 120 days of an annual OSCAR data point, resulting in 399,206 resident-level observations.
Principal Findings. Controlling for endogeneity of staffing increases the estimated impact of staffing on outcomes in nursing homes. Greater RN staffing significantly decreases the likelihood of both adverse outcomes. Increasing skill mix only reduces the incidence of urinary tract infections.
Conclusions. Research that fails to account for endogeneity of the staffing–outcomes relationship may underestimate the benefit from increased RN staffing. Increases in RN staffing are likely to reduce adverse outcomes in some nursing homes. More research using a broader array of instruments and a national sample would be beneficial.

Key Words. Staffing, nursing homes, quality of care, instrumental variables, endogeneity

The quality of care in nursing homes has long been an issue of concern to policy makers and the public. Following the Institute of Medicine report that documented serious abuses and quality of care problems (IOM 1986), intense scrutiny and extensive regulations have led to substantial improvements in the quality of nursing home care. However, many challenges remain, as documented by the persistently high rates of deficiency citations for problems such as pressure sores, food sanitation, unnecessary drugs, and infection control (Harrington and Carrillo 1999). More recent calls to improve quality have focused on the need to improve staffing levels, particularly registered nurse (RN) staffing. A government-sponsored report found that the vast majority of nursing facilities do not provide sufficient staffing to ensure basic quality (Abt Associates 2001), and a group of expert researchers independently called for higher nurse staffing requirements (Harrington, Kovner et al. 2000).

While increasing staffing to improve quality seems like common sense, little solid evidence exists upon which to base specific recommendations or justify the cost of increased staffing. Whether and to what extent increased staffing will improve outcomes is uncertain, as the current understanding is based largely on cross-sectional evidence prone to omitted variable and endogeneity bias. This study examines the staffing–outcomes relationship in nursing homes with the intention of bridging that gap in the evidence.

BACKGROUND

Considerable research has been devoted to the issues of the number and composition of nursing staff required to meet the needs of nursing home residents. Not surprisingly, most findings have suggested that a higher nursing staff intensity (i.e., more care hours per resident-day) and more skilled nursing staff mix (i.e., a greater proportion of professional nursing staff such as RNs) are associated with higher-quality care as measured by various process and outcome indicators. Examples include improved survival (Cohen and Spector 1996; Porell et al. 1998), better functional status (Cohen and Spector 1996), less incontinence (Porell et al. 1998), fewer pressure sores (Cohen and Spector 1996; Weech-Maldonado et al. 2004), lower rates of hospitalization (Carter and Porell 2003) and physical restraint use (Castle 2000; Weech-Maldonado et al. 2004), and fewer facility deficiencies (Harrington, Zimmerman et al. 2000).

While nurse aides (NAs) provide the majority of direct care to nursing home residents, efforts to improve quality have focused more on increasing professional nursing staff. The care provided by NAs is nontechnical and consists primarily of helping residents with activities of daily living (ADLs)

Address correspondence to R. Tamara Konetzka, Ph.D., Assistant Professor, University of Chicago, Department of Health Studies, 5841 S. Maryland Avenue, MC2007, Chicago, IL 60637. Sally C. Stearns, Ph.D., Associate Professor, is with the University of North Carolina at Chapel Hill, Department of Health Policy and Administration, Chapel Hill, NC. Jeongyoung Park, Ph.D., Post-Doctoral Fellow, is with the University of Pennsylvania, Division of General Internal Medicine, Guardian Drive, Philadelphia, PA.

such as eating, dressing, bathing, toileting, and walking (Cawley, Grabowski, and Hirth 2006). When NAs provide direct care, RNs observe, assess, and record resident symptoms and progress. RNs also collaborate with physicians in treatment, administration of medications, and development of care plans.

Existing evidence supports that RN hours may be more important than total nursing hours per resident. In a study using a nationally representative sample of nursing homes and residents, Cohen and Spector (1996) found that a higher RN intensity was associated with a lower rate of mortality; however, having more NAs had no impact on resident outcomes. Horn et al. (2005) found that more RN hours were associated with fewer pressure ulcers, hospitalizations, and urinary tract infections (UTIs); less weight loss, catheterization, and deterioration in the ability to perform ADLs; and greater use of nutritional supplements, while more NA hours were associated only with fewer pressure sores.

Using Online Survey, Certification and Reporting (OSCAR) data from 1992 to 1997, Castle (2000) found that facilities with more RNs per 100 beds were less likely to increase restraint use. Harrington, Zimmerman et al. (2000) examined the relationship between staffing hours per resident-day and facility deficiencies identified by state surveyors under federal certification regulations. Consistent with previous research, fewer RN hours was associated with more quality of care deficiencies. More recently, Decker (2006) used National Nursing Home Survey Data from 1999 to show that RN staffing levels were associated with faster discharge to the community for short-stay residents requiring primarily postacute care but not for long-stay residents.

Due to the difficulty in finding adequate numbers of qualified staff or the desire to minimize costs, some facilities may hire fewer than the optimal number of RNs or may shift some tasks typically done by RNs to less qualified nursing staff, which may have implications for resident outcomes. Using a sample of 1,287 nursing homes in five states, Weech-Maldonado et al. (2004) found that greater use of RNs, both in absolute terms and relative to total nursing staff, was associated with a reduced likelihood of pressure ulcers, better cognitive functioning, and lower use of restraints. This finding suggests that staff mix may be as important as the level, and that efforts to improve quality should focus on increasing the proportion of professional staff as well as RN staff intensity.

In summary, most evidence indicates that more RN staffing, both in an absolute sense and as a proportion of total hours, is associated with better outcomes. The cross-sectional research, however, is subject to omitted variable bias, and most studies except Cohen and Spector (1996) are subject to possible bias from reverse causality. Therefore, mandating and enforcing higher staffing ratios could have no effect on outcomes. Little longitudinal evidence is available. Zhang and Grabowski (2004) used longitudinal OSCAR data to look at whether the Nursing Home Reform Act improved both staffing and quality. Implementation of this act was shown to increase staffing, but the staffing increases were not associated with improvements in quality except in homes that were of particularly poor quality at baseline. The facility-level measures of quality available in OSCAR may be insufficiently sensitive, and endogeneity of staffing and outcomes was not addressed. The use of a longitudinal model in a recent study of the effect of hospital staffing on mortality was shown to result in substantially smaller estimates than those from cross-sectional research (Mark et al. 2004), with a larger RN effect when endogeneity was addressed.

We therefore model resident-level outcomes using a longitudinal model with facility fixed effects and instrumental variables to correct for endogeneity. The fixed-effects model, which controls for any omitted time-invariant facility-level variables, gives us the effect of a *change* in RN staffing and skill mix on a *change* in outcomes. While the fixed-effects model is an improvement over cross-sectional analyses, it does not control for any omitted time-varying variables or address all potential endogeneity of staffing and outcomes. We view the staffing decision as inherently endogenous. That is, facility managers make structural decisions about staffing (both intensity and skill mix) and the quality of care to be provided subject to regulations, resource constraints, and current and expected case-mix. At the same time, these decisions affect resident outcomes. Because facilities with sicker residents would tend to opt for higher staffing, the endogeneity bias would tend to underestimate the effect of staffing on outcomes. We therefore expect that accounting for this endogeneity through instrumental variables will result in estimates of larger magnitude than in a longitudinal model without adjustment for endogeneity.

METHODS

Data

We analyzed clinical outcomes and case-mix data from the Nursing Facility Minimum Data Set (MDS), a government-mandated data set containing assessment data on all residents in Medicare- or Medicaid-certified nursing facilities. Our analysis used quarterly MDS assessments (excluding those with stays <90 days) from all freestanding nursing homes in Ohio, Kansas, Maine, Mississippi, and South Dakota that existed from 1997 to 2000. We chose these states because they had reliable data before and after implementation of the policy changes that we use as instruments and because they provided a large sample of facilities and residents from various geographic regions. We excluded hospital-based facilities as they are fundamentally different than freestanding facilities in terms of staffing and do not have the long-stay populations that are the focus of this study. The sample represents approximately 10 percent of freestanding facilities and residents nationwide. A limitation of this analysis is generalizability from the five-state sample to the United States; however, characteristics of nursing homes in the five states do not appear to diverge widely from national norms.

We combined MDS data with facility-level information from the Centers for Medicare and Medicaid Services' (CMS) OSCAR database. Staffing ratios and all facility-level controls were taken from OSCAR. We followed standard procedures to calculate staffing ratios, assuming that each full-time equivalent staff member works 70 hours in a 2-week period and dividing the staffing hours per day by the number of residents in the facility (Abt Associates 2001). Although OSCAR is the only uniform and easily available source for nurse staffing data, it has several limitations. The certification procedures are generally not audited, which raises concerns regarding validity and reliability of the data. Furthermore, staffing is reported for a 2-week period at the time of survey, so it may not accurately depict the facility's staffing over a longer period. In particular, it may overstate staffing if the facility increases staffing during the period around the survey (Harrington, Kovner et al. 2000; Harrington, Zimmerman et al. 2000; Zhang and Grabowski 2004). Nonetheless, OSCAR data have been widely used for nursing home studies. Calculation of our instruments was from both OSCAR data and Medicare Cost Reports from CMS. Construction of the instruments is described in the instrumental variables section below.

Some quarterly MDS assessments were not available or were excluded from the analysis. No assessments were available for the first half of 1998, resulting in a uniformly smaller sample for that year. Because OSCAR data are collected only once per year on average (resulting in just one staffing data point per year) but MDS data are collected on an ongoing basis, we included only assessments that were conducted within 120 days before or after an OSCAR survey. (Experimentation with different ranges showed that relationships were attenuated with a wider range and power was substantially decreased with a smaller range.) The final sample included 399,206 resident assessments from 1,366 facilities.

Analysis

We estimate facility-level fixed-effects models of the effect of RN staffing and skill mix on two resident outcome measures: pressure sores and UTIs.

We implement fixed effects with a dichotomous outcome using conditional logit regression in *STATA*. The fixed effects conditional logit models control for all time-invariant variables at the facility, market, or state level, such as proprietary status, general managerial approach and skills, urban/rural location, state policies, and affluence of the area.

Outcomes are modeled as a function of RN staffing intensity, skill mix, resident-level severity controls, time-varying facility- and market-level controls, facility-level fixed effects and time fixed effects. The basic model for resident i in facility f at time t thus has the following form:

$$\begin{aligned}\Pr(\mathit{Outcome})_{ift} = {} & \beta_0 + \beta_1 \mathit{RN}_{ft} + \beta_2 \mathit{SkillMix}_{ft} \\ & + \beta_3 \mathit{Severity}_{ift} + \beta_4 \mathit{Facility}_{ft} + \beta_5 \mathit{Year}_t\end{aligned}$$

where *Severity* is a vector of resident-level controls, *Facility* is a vector of time-varying facility- and market-level controls, and *Year* represents time fixed effects. The time fixed effects are three indicator variables for years 1998–2000 that account for any underlying time trend. State-year interactions were tested and found not to affect results in a meaningful way, so they were not included. We estimated the models with and without instrumental variables techniques to account for the endogeneity of the staffing-outcomes relationship. All models also adjusted the standard errors for clustering due to repeated observations on individual residents.

The two outcome measures used in this analysis were chosen from the literature to reflect important and likely staffing-dependent aspects of care affecting long-stay residents and because they occur frequently enough to be sensitive to financial pressures. Recently, CMS judged pressure sores and urinary tract infections to be among the chronic-care quality indicators with the highest validity (Morris et al. 2003). Prevention of pressure sores and UTIs involves frequent position change, proper hydration, and careful hygiene, which are primarily the responsibilities of NAs, but subject to proper supervision and detection by RNs. Each outcome, represented by a binary variable equal to 1 if the resident experienced the event at the time of the assessment and 0 otherwise, was analyzed at the resident level to allow for the most precise resident-specific risk adjustment. Only stage 2 and above pressure ulcers were used in calculating the pressure sore measure, as stage 1 ulcers may be especially prone to ascertainment bias.

Staffing was measured with two variables. RN staffing intensity was calculated as RN hours per resident-day. Skill mix was measured as RN staffing hours as a proportion of total staffing hours (RN, LPN, and NA combined).

While the two measures are related, RN hours per resident day measures the extent of RN expertise available for resident care, while the skill mix measure represents the extent to which greater supervision by RNs may be available or pressure to shift tasks to less qualified staff may be reduced. We did not include separate variables for LPN and NA hours because these become redundant with the RN and skill mix combination; one cannot interpret a change in skill mix while holding all types of staffing hours constant.

Resident-level severity controls include age, gender, a set of diagnosis indicators, dependence in ADLs, and the Cognitive Performance Score (Morris et al. 1994). Because ADL and cognitive functioning may themselves be outcome measures, we use lagged values of these controls. A lagged value of each dependent variable is also included, as residents who once had a pressure sore or UTI may be more susceptible to them. The lagged values are based on the assessment closest to the previous OSCAR staffing data point. We also control for Medicare-covered stays, as Medicare residents are often rehabilitation patients who may be different in unmeasured ways from more typical long-stay residents. Payer source data in the MDS appear to contain frequent errors and inconsistencies; these data were cleaned to the extent possible but likely contain some measurement error.

Facility-level case-mix may affect outcomes above and beyond individual severity controls, as facilities with higher case mix may be more skilled at serving residents with higher need. Facility-level case-mix controls in the analysis include: a measure of ADL functioning and an index of skilled services provided (e.g., tubefeeding, IV drugs, ventilator care) (Cowles 2002); the percent of residents in the facility with dementia, depression, psychiatric diagnoses; the percent whose care is paid for by Medicare, which may also measure resources available to the facility (Konetzka, Spector, and Shaffer 2004; Konetzka, Yi et al. 2004; Konetzka, Norton et al. 2006); and the percent private-pay and facility occupancy rate, which may also indicate resource availability. To account for market competitiveness, we include a Herfindahl–Hirschman index (HHI), calculated as the sum of squared market shares of all nursing homes in each county. Summary statistics on dependent and explanatory variables are shown in Table 1.

Instrumental Variables Analysis

Both staffing variables in the analysis, RN staffing hours per resident-day and skill mix, are potentially endogenous. Instrumental variables are designed to address endogeneity and produce less biased estimates. First stage equations

Table 1: Summary Statistics on the Sample of Nursing Home Resident Assessments from Five States, $n = 399{,}206$

Variable	*Mean or Proportion*	*Standard Deviation*
Outcome variables		
Pressure sore within last 14 days	0.043	
UTI within last 30 days	0.067	
Staffing variables		
RN hours per resident-day	0.350	0.219
Skill mix (% of total hours provided by RNs)	0.117	0.064
Time trend indicators		
1997	0.225	
1998	0.104	
1999	0.311	
2000	0.360	
Resident-level characteristics		
Age	82.5	12.0
Female	0.759	
Medicare payment source (yes = 1)	0.032	
Lagged pressure sore	0.042	
Lagged UTI	0.071	
Lagged ADL score	10.4	5.0
Lagged cognitive performance score	2.89	1.81
Comatose	0.004	
Alzheimers/other dementia	0.711	
Stroke	0.237	
Heart disease	0.382	
Cancer	0.075	
Diabetes	0.215	
Depression	0.407	
Facility- and market-level characteristics		
For-profit facility	0.698	
Nonprofit facility	0.256	
Government-owned facility	0.046	
Chain-owned facility	0.546	
Proportion of residents with Medicare payer source	0.063	0.064
Proportion of residents with private payer source	0.268	0.179
Facility occupancy rate	0.867	0.129
ADL index (range 3.2–15.8)	10.2	1.3
Special care index	0.170	0.119
Percent of residents with dementia	0.499	0.158
Percent of residents with depression	0.443	0.217
Percent of residents with psychiatric disorders	0.198	0.178
Herfindahl–Hirschman index (HHI)	0.232	0.237

UTI, urinary tract infection; ADL, activities of daily living.

for the dependent staffing variables (RN hours per resident-day and skill mix) are estimated as a function of the instruments, other explanatory variables, and facility-level fixed effects using facility-year observations. In this analysis, a valid instrument must predict RN staffing and/or skill mix but not affect the incidence of pressure sores and UTIs other than through staffing.

We use an important Medicare policy change, the introduction of a Prospective Payment System (PPS) for Medicare payment in nursing homes, as an exogenous financial shock to nursing facilities that affected RN staffing. This new policy dramatically changed the form of payment for Medicare services in skilled nursing facilities (SNFs) by replacing the former cost-based reimbursement system with a PPS. At the same time, the overall level of funding was reduced dramatically, reducing the average Medicare reimbursement for the majority of SNFs. Several recent studies provide evidence that this financial shock to nursing homes resulted in decreased RN staffing proportional to each facility's dependence on Medicare (Konetzka, Yi et al. 2004; Konetzka, Norton et al. 2006; White 2005). One study used additional variation provided by staggered implementation of the new system, finding that the timing of implementation could be tied to reductions in staffing (Konetzka, Yi et al. 2004). Thus, PPS is a compelling predictor of RN staffing ratios within facilities over time. These studies asserted that PPS affected outcomes mainly through staffing, but did not test this assertion in part because of the endogeneity issues explored here.

We use a dichotomous indicator of PPS specific to each facility as the primary instrument. The indicator is equal to 1 for all assessments under the new PPS system, which was implemented beginning with each facility's fiscal year start date on or after July 1, 1998. Thus, our instrument has variation among facilities in the timing of implementation. For our second instrument, we multiply PPS by the percent of residents in the facility with Medicare payer source at baseline (1997) to capture the extent of the shock; facilities with more Medicare should have experienced a greater shock to financial resources and thus to staffing ratios. Percent Medicare is defined as the percent of total residents in a facility whose primary payer is Medicare. Medicare cost reports for 1997 were used to define the percent of resident-days payable by Medicare; mean OSCAR values for baseline years were used for facilities without Medicare cost reports because the OSCAR values may fluctuate somewhat from year to year. The combination of these two data sets provided stable estimates of the percent Medicare.

The financial shock of PPS may have affected pressure sores and UTIs through mechanisms other than nurse staffing; for example, through decreased capital expenditures. This would be a violation of the assumptions of the instrumental variables model. We argue that any nonstaffing effects of PPS on the outcomes we study are likely to be very small for three reasons: (1) The majority of a nursing facility's budget is staffing; (2) The outcomes we study are likely to be particularly staffing-sensitive; and (3) Any changes to capital expenditures resulting from PPS are likely to be longer-run changes, while high turnover rates enable almost immediate adjustments to staffing ratios as demonstrated in previous research (Konetzka, Yi et al. 2004). Because our system is exactly identified (two right hand side potentially endogenous variables and two instruments), we cannot explicitly test this assumption, but we conduct a sensitivity analysis to measure the extent of the potential bias under varying assumptions. Details and results of the sensitivity analysis can be found in Appendix B.

The instrumental variables analysis in this paper is conducted using two alternative methods, a traditional two-stage least squares (2SLS) and a residual-inclusion approach (Blundell and Smith 1989, 1993; Terza, Basu, and Rathouz forthcoming). For the traditional 2SLS, we obtain predicted values of the staffing measures from the first stage fixed-effects OLS regression, and use those predicted values in the outcome equation. The problem with the 2SLS approach is that it assumes a linear model for the second stage estimation; its application to a nonlinear model may result in inconsistent parameter estimates. Blundell and Smith and (separately) Terza have shown that consistent parameter estimates can be obtained by including the predicted residuals from the first stage estimation in the second equation (i.e., the outcome equation of interest estimated using a conditional logit approach).

In principle, because these methods include a predicted value in the main equation, standard errors should be adjusted. As standard methods of calculating this adjustment do not apply to conditional logit estimation, we do not correct the standard errors for the two-stage approach; however, we believe the effects of such a correction would be minimal and would not qualitatively change our results or conclusions.

Marginal Effects

In order to determine the clinical significance of the results, it is useful to calculate marginal effects of the staffing variables on outcomes. The precise calculation of the marginal effects is complicated, however, by the nonlinear nature of the conditional logit model and by the use of two variables that both involve RN time (i.e., so that the calculation of the effect of a change in RN hours at a facility, holding all else constant, would require joint consideration of both coefficients). We use a simple approximation to marginal effects generally used in logit models, calculated separately for RN staffing and skill mix at the mean outcome level of the sample. An alternative specification used linear probability models (LPMs) with joint consideration of RN staffing and skill mix; results were similar in magnitude and direction.

RESULTS

Stage 1 and Specification Tests

A joint F-test on the two instruments (PPS and PPS $\times$ Percent Medicare) in the first-stage equation confirms that the instruments are good predictors of RN staffing ($F = 19$) and Skill Mix ($F = 18$), even within the context of a fixed-effects model where strong instruments are difficult to find. In addition, both instruments are individually significant ($p < .01$ in all cases except PPS in the RN regression with $p < .05$). See Appendix A for details of the first-stage regressions. A Durbin–Wu–Hausman test (Hausman 1978) was conducted for each potentially endogenous variable to determine whether exogeneity could be rejected for each of the hypothesized endogenous variables. Exogeneity could not be rejected in the case of skill mix in the pressure sore outcome equation; we therefore correct skill mix for endogeneity only in the UTI regressions.

Stage 2 Main Equation

Tables 2 and 3 present the model estimation results for pressure sores and UTIs, respectively. Each table provides results for the three models (conditional logit, conditional logit with traditional IV, and conditional logit with residual inclusion). The conditional logit without IV produces small coefficients that are marginally significant at conventional levels for the effects of RN staff intensity, and the skill mix measure is not statistically significant. The sign on the RN staffing-level variable differs, however, across the two outcome equations. The likelihood of a resident having a pressure sore decreases with greater RN intensity, while the likelihood of a resident having a UTI increases with greater RN intensity. Both coefficients may be biased upward because of unmeasured case-mix heterogeneity that can be addressed with the IV approach.

The traditional IV and residual inclusion IV coefficient estimates are very similar to each other, but very different from the estimation without IV. We focus on the IV conducted with the residual inclusion approach, as this

Table 2: Effects of Staffing on Pressure Sores (Conditional Logit Models)

	Facility Fixed Effects Model	*Facility Fixed Effects with Traditional IV*	*Facility Fixed Effects with Residual Inclusion IV*
Staffing variables			
RN hours per resident day	– 0.222* (0.123)	– 3.006*** (0.515)	– 3.002*** (0.515)
Skill mix	0.632 (0.424)	– 0.009 (0.254)	0.045 (0.437)
RN residual			2.983*** (0.536)
Time trends			
1998	0.008 (0.033)	– 0.005 (0.033)	– 0.005 (0.033)
1999	0.002 (0.027)	– 0.007 (0.027)	– 0.007 (0.027)
2000	0.006 (0.027)	– 0.033 (0.028)	– 0.032 (0.028)
Resident-level controls			
Age spline 1	0.769*** (0.187)	0.775*** (0.187)	0.775*** (0.187)
Age spline 2	– 0.134*** (0.047)	– 0.134*** (0.047)	– 0.134*** (0.047)
Age spline 3	– 0.040* (0.024)	– 0.041* (0.024)	– 0.041* (0.024)
Age spline 4	0.064*** (0.016)	0.066*** (0.016)	0.066*** (0.016)
Age spline 5	0.032 (0.071)	0.028 (0.071)	0.028 (0.071)
Female	– 0.235*** (0.020)	– 0.237*** (0.020)	– 0.237*** (0.020)
Medicare payer source	0.796*** (0.039)	0.806*** (0.039)	0.806*** (0.039)
Lagged pressure sore	1.655*** (0.022)	1.654*** (0.022)	1.655*** (0.022)
Lagged ADL score	0.155*** (0.002)	0.156*** (0.002)	0.156*** (0.002)
Lagged cognitive performance	– 0.050*** (0.005)	– 0.051*** (0.005)	– 0.051*** (0.005)
Comatose	0.518*** (0.089)	0.514*** (0.089)	0.514*** (0.089)
Alzheimers/other dementia	– 0.045 (0.038)	– 0.040 (0.038)	– 0.040 (0.038)
Stroke	– 0.091*** (0.019)	– 0.090*** (0.019)	– 0.090*** (0.019)
Heart disease	0.117*** (0.018)	0.113*** (0.018)	0.113*** (0.018)
Cancer	0.050 (0.032)	0.048 (0.032)	0.048 (0.032)
Diabetes	0.358*** (0.019)	0.359*** (0.019)	0.359*** (0.019)
Depression	– 0.058*** (0.018)	– 0.060*** (0.018)	– 0.060*** (0.018)
Facility-level controls			
Medicare resident fraction	0.157 (0.261)	0.1614 (0.261)	0.162 (0.262)
Private pay fraction	– 0.001 (0.002)	– 0.001 (0.002)	– 0.001 (0.002)
Occupancy rate	– 0.060 (0.174)	– 0.040 (0.173)	– 0.044 (0.174)
ADL index	0.012 (0.015)	0.001 (0.015)	0.001 (0.015)
Special care index	0.085 (0.171)	0.081 (0.171)	0.081 (0.171)
Percent dementia	0.123 (0.103)	0.314*** (0.109)	0.314*** (0.109)
Percent depressed	0.064 (0.070)	0.127* (0.071)	0.126* (0.071)
Percent psychiatric	– 0.106 (0.095)	– 0.079 (0.096)	– 0.079 (0.096)
HHI	0.090 (0.328)	0.583 (0.341)	0.583* (0.341)
Number of observations	393,178	393,178	393,178

Standard errors in parentheses.

*Significant at 10%.

***Significant at 1%.

ADL, activities of daily living; HHI, Herfindahl–Hirschman index.

Table 3: Effects of Staffing on UTIs (Conditional Logit Models)

	Facility Fixed Effects Model	*Facility Fixed Effects with Traditional IV*	*Facility Fixed Effects with Residual Inclusion IV*
Staffing variables			
RN hours per resident day	0.194* (0.106)	– 1.528*** (0.410)	– 1.556*** (0.411)
Skill mix	– 0.504 (0.352)	– 1.634*** (0.525)	– 1.661*** (0.495)
RN residual			1.683*** (0.429)
Skill mix residual			1.544** (0.665)
Time trends			
1998	– 0.013 (0.026)	– 0.021 (0.026)	– 0.020 (0.026)
1999	– 0.018 (0.021)	– 0.020 (0.021)	– 0.019 (0.021)
2000	– 0.033 (0.021)	– 0.056*** (0.022)	– 0.055*** (0.022)
Resident-level controls			
Age spline 1	0.280* (0.166)	0.285* (0.167)	0.285* (0.166)
Age spline 2	– 0.167*** (0.043)	– 0.167*** (0.043)	– 0.166*** (0.043)
Age spline 3	0.001 (0.020)	0.001 (0.020)	0.001 (0.020)
Age spline 4	0.020 (0.012)	0.021* (0.012)	0.021* (0.012)
Age spline 5	– 0.161*** (0.060)	– 0.164*** (0.060)	– 0.164*** (0.060)
Female	0.250*** (0.017)	0.249*** (0.017)	0.249*** (0.017)
Medicare payer source	0.576*** (0.034)	0.582*** (0.034)	0.582*** (0.034)
Lagged UTI	1.073*** (0.017)	1.077*** (0.017)	1.077*** (0.017)
Lagged ADL score	0.063*** (0.002)	0.063*** (0.002)	0.063*** (0.002)
Lagged cognitive performance	– 0.061*** (0.004)	– 0.062*** (0.004)	– 0.062*** (0.004)
Comatose	– 0.016 (0.105)	– 0.021 (0.105)	– 0.019 (0.105)
Alzheimers/other dementia	0.010 (0.030)	0.013 (0.030)	0.013 (0.030)
Stroke	0.048*** (0.015)	0.049*** (0.015)	0.049*** (0.015)
Heart Disease	0.038*** (0.014)	0.036** (0.014)	0.036** (0.014)
Cancer	0.143*** (0.024)	0.142*** (0.024)	0.142*** (0.024)
Diabetes	0.221*** (0.015)	0.221*** (0.015)	0.221*** (0.015)
Depression	0.135*** (0.014)	0.134*** (0.014)	0.134*** (0.014)
Facility-level controls			
Medicare resident fraction	0.073 (0.200)	0.094 (0.200)	0.087 (0.200)
Private pay fraction	– 0.001 (0.001)	– 0.001 (0.001)	– 0.001 (0.001)
Occupancy rate	0.010 (0.137)	0.019 (0.137)	0.041 (0.137)
ADL index	– 0.009 (0.012)	– 0.014 (0.012)	– 0.015 (0.012)
Special care index	0.196 (0.139)	0.179 (0.139)	0.186 (0.139)
Percent dementia	0.131 (0.081)	0.251*** (0.085)	0.253*** (0.085)
Percent depressed	0.118** (0.055)	0.154*** (0.055)	0.157*** (0.055)
Percent psychiatric	0.092 (0.074)	0.115 (0.074)	0.117 (0.074)
HHI	– 0.602*** (0.229)	– 0.285 (0.238)	– 0.293 (0.238)
Number of observations	399,206	399,206	399,206

Standard errors in parentheses.

*Significant at 10%.

**Significant at 5%.

***Significant at 1%.

UTI, urinary tract infection; ADL, activities of daily living; HHI, Herfindahl–Hirschman index.

method should yield consistent parameter estimates. For pressure sores (Table 2), greater levels of RN intensity are associated with significantly lower rates of pressure sores, with an estimated coefficient that is more than 13 times greater than the non-IV estimates (-3.002 versus -0.222). The skill mix variable is not statistically significant in any of the pressure sore estimations. For UTI (Table 3), greater RN intensity as well as a more highly skilled staffing mix result in negative and statistically significant coefficients, indicating better outcomes.

The marginal effects are considered in two ways. First, we use standard approximations to average marginal effects in logit models to consider the predicted levels of outcomes for a 1-unit increase in RN intensity and in skill mix, calculated separately (i.e., assuming that the other variable remains constant). The estimates are quite large. The incidence of pressure sores would be almost eliminated by the increase in RN staffing and the incidence of UTIs would be reduced from 6.7 percent to approximately 1.5 percent. However, it is crucial to note that a 1-unit change in RN staffing or in skill mix represents an unrealistically large increment, and that it is not possible to hold skill mix constant when increasing RN staffing unless LPN and NA staffing are also increased. The mean of RN hours per resident-day is 0.35, so a 1-unit increase would mean almost a quadrupling of RN time to 1.35 hours per resident-day. The mean of skill mix in the sample is 0.117, and a 1-unit change would effectively imply moving from almost no RNs to all RNs.

A more realistic change that allows both staffing measures to vary might be a 50 percent increase in RN time to 0.525 hours per resident-day, which also means an increase in skill mix to 0.163 on average in the sample, holding other staffing levels constant. The effect of these smaller increases on outcomes cannot be calculated using standard logit approximations, but can be estimated with LPMs. While we cannot assume that the marginal effects follow a linear pattern, we found it useful to estimate a ballpark marginal effect for the 50 percent increase using an LPM and incorporating the joint effects of an increase in RN staffing and skill mix. (The base results from the LPM were very similar to the conditional logit models.) For a 50 percent increase in RN hours per resident-day, the rate of pressure sores is predicted to decline by about 66 percent and that of UTIs by about 45 percent for the average facility.

Sensitivity Analysis

Our results appear to be robust to reasonable violations of the exclusion restriction, but our magnitudes of effect may be overestimated. Please see Appendix B for details.

DISCUSSION

In this paper we analyzed the effects of RN staffing and skill mix on several important outcomes in nursing homes. While most prior evidence on this topic stems from cross-sectional research subject to substantial omitted variable bias, our analysis uses a longitudinal fixed-effects design that estimates effects based on changes in staffing and outcomes over time. In addition, we employ instrumental variables to address the potential endogeneity of the staffing-outcomes relationship.

We find that increased RN staffing improves outcomes and that the magnitude of the effect is statistically and practically significant. The effects of skill mix, however, vary by outcome. This result is consistent with the type of care leading to each outcome (Horn et al. 2005). Avoidance of pressure sores is more dependent on NAs for turning, and a high skill mix may mean that fewer aides are available, potentially balancing out the benefits of increased RN expertise. Avoidance of UTIs, however, may be dependent on a higher level of skill, e.g., for proper catheter care. In any case, the effects of skill mix appear to be small compared with effects of the absolute number of RN hours. Finally, we find that the staffing-outcomes relationship is likely to be endogenous, and that failure to correct for this endogeneity may lead to underestimates of the benefits from increased staffing.

As with most instrumental variables analyses, some caution should be used in interpreting these results. First, it is important to note that the IV estimates are of local average treatment effects and therefore may not apply to all facilities (Harris and Remler 1998; Newhouse and McClellan 1998). In this case, the estimates are identified off facilities that decreased case-mix-adjusted staffing in response to the financial shock of PPS but would not have decreased staffing in the absence of this financial shock. Facilities that decreased staffing (or failed to increase staffing as case-mix increased) may not be representative of all nursing homes, and perhaps more importantly, the magnitude of effect on outcomes may not be the same for a commensurate increase in staffing.

Second, while we believe our PPS instruments are strong, we cannot conduct tests of over-identification that would support their exclusion from the main equation. We argue that PPS affected the incidence of pressure sores and UTIs among long-stay populations through changes in staffing, and that any other effects of PPS on these outcomes were likely to be very small. Based on the results of our sensitivity analysis, the effects through other pathways would have to be quite large to invalidate our conclusions. Additional studies using a nationwide sample, different instruments, or employing other methods to reduce omitted variable and endogeneity bias (e.g., randomization), would be beneficial.

We note that our sample was limited to long-stay residents; the results may or may not be applicable to post-acute care populations in nursing homes, though one could speculate that similar endogeneity issues apply. It is also important to note that this study is unable to test the mechanisms by which the outcomes may be affected by increased staffing. While we hypothesize that greater supervision and skill in ascertainment of problems by RNs would lead to changes in our selected outcomes, we are unable to substantiate those processes. Additional research collecting detailed process data to fill in the mechanisms of the staffing–outcomes relationship is needed in order to strengthen the case for causality.

The need for higher staffing in nursing homes, especially RN staffing, has received a great deal of attention in recent years. Many states have, or are considering, legislation that mandates minimum RN staffing ratios. However, the research on the staffing–outcomes relationship is still highly inadequate to inform such policy decisions. Cross-sectional research is likely to produce incorrect estimates of the magnitude and significance of effect due to omitted variable bias, and longitudinal research may still produce incorrect estimates due to endogeneity. This study finds that methodological approach matters and that failure to account for endogeneity of staffing decisions may result in an underestimate of the benefit from RN staffing in nursing homes.

ACKNOWLEDGMENTS

The authors would like to thank John Nyman, Anirban Basu, participants in Northwestern University's Buehler Center on Aging workshop, and two anonymous reviewers for helpful comments and advice. MDS data were acquired with funds from Beverly Enterprises Inc. and were used under Data Use Agreement 11521 with CMS.

REFERENCES

Abt Associates. 2001. *Report to Congress: Appropriateness of Minimum Nurse Staffing Ratios in Nursing Homes Phase II Final Report.* Baltimore, MD: Centers for Medicare and Medicaid Services.

Blundell, R. W., and R. J. Smith. 1989. "Estimation in a Class of Simultaneous Equation Limited Dependent Variable Models." *Review of Economics and Statistics* 56: 37–58.

———. 1993. "Simultaneous Microeconmetric Models with Censored or Qualitative Dependent Variables." In *Handbook of Statistics*, Vol 2, edited by G. S. Maddala, C. R. Rao, and H. D. Vinod, pp. 1117–43. Amsterdam: North Holland Publishers.

Carter, M. W., and F. W. Porell. 2003. "Variations in Hospitalization Rates among Nursing Home Residents: The Role of Facility and Market Attributes." *Gerontologist* 43 (2): 175–91.

Castle, N. G. 2000. "Differences in Nursing Homes with Increasing and Decreasing Use of Physical Restraints." *Medical Care* 38 (12): 1154–63.

Cawley, J., D. C. Grabowski, and R. A. Hirth. 2006. "Factor Substitution in Nursing Homes." *Journal of Health Economics* 25 (2): 234–47.

Cohen, J. W., and W. D. Spector. 1996. "The Effect of Medicaid Reimbursement on Quality of Care in Nursing Homes." *Journal of Health Economics* 15 (1): 23–48.

Cowles, C. 2002. *2002 Nursing Home Statistical Yearbook*. Montgomery Village, MD: Cowles Research Group.

Decker, F. H. 2006. "Nursing Staff and the Outcomes of Nursing Home Stays." *Medical Care* 44 (9): 812–21.

Harrington, C., and H. Carrillo. 1999. "The Regulation and Enforcement of Federal Nursing Home Standards, 1991–1997." *Medical Care Research and Review* 56 (4): 471–94.

Harrington, C., C. Kovner, M. Mezey, J. Kayser-Jones, S. Burger, M. Mohler, R. Burke, and D. Zimmerman. 2000. "Experts Recommend Minimum Nurse Staffing Standards for Nursing Facilities in the United States." *Gerontologist* 40 (1): 5–16.

Harrington, C., D. Zimmerman, S. L. Karon, J. Robinson, and P. Beutel. 2000. "Nursing Home Staffing and Its Relationship to Deficiencies." *Journal of Gerontology Series B: Psychological Science and Social Science* 55 (5): S278–87.

Harris, K. M., and D. K. Remler. 1998. "Who Is the Marginal Patient? Understanding Instrumental Variables Estimates of Treatment Effects." *Health Services Research* 33 (5, part 1): 1337–60.

Hausman, J. A. 1978. "Specification Tests in Econometrics." *Econometrica* 46 (6): 1251–71.

Horn, S. D., P. Buerhaus, N. Bergstrom, and R. J. Smout. 2005. "RN Staffing Time and Outcomes of Long-Stay Nursing Home Residents: Pressure Ulcers and Other Adverse Outcomes Are Less Likely as RNs Spend More Time on Direct Patient Care." *American Journal of Nursing* 105 (11): 58–70; quiz 71.

Institute of Medicine (IOM), Committee on Nursing Home Reform. 1986. *Improving the Quality of Care in Nursing Homes*. Washington, DC: National Academy Press.

Konetzka, R. T., E. C. Norton, P. D. Sloane, K. E. Kilpatrick, and S. C. Stearns. 2006. "Medicare Prospective Payment and Quality of Care for Long-Stay Nursing Facility Residents." *Medical Care* 44 (3): 270–6.

Konetzka, R. T., W. Spector, and T. Shaffer. 2004. "Effects of Nursing Home Ownership Type and Resident Payer Source on Hospitalization for Suspected Pneumonia." *Medical Care* 42 (10): 1001–8.

Konetzka, R. T., D. Yi, E. C. Norton, and K. E. Kilpatrick. 2004. "Effects of Medicare Payment Changes on Nursing Home Staffing and Deficiencies." *Health Services Research* 39 (3): 463–88.

Mark, B. A., D. W. Harless, M. McCue, and Y. Xu. 2004. "A Longitudinal Examination of Hospital Registered Nurse Staffing and Quality of Care." *Health Services Research* 39 (2): 279–300.

Morris, J., T. Moore, R. Jones, V. Mor, J. Angelelli, K. Berg, C. Hale, S. Morris, K. Murphy, and M. Rennison. 2003. *Validation of Long-Term and Post-Acute Care Quality Indicators.* Baltimore: Centers for Medicare and Medicaid Services (CMS).

Morris, J. N., B. E. Fries, D. R. Mehr, C. Hawes, C. Phillips, V. Mor, and L. A. Lipsitz. 1994. "MDS Cognitive Performance Scale." *Journals of Gerontology* 49 (4): M174–82.

Newhouse, J. P., and M. McClellan. 1998. "Econometrics in Outcomes Research: The Use of Instrumental Variables." *Annual Review of Public Health* 19: 17–34.

Porell, F., F. G. Caro, A. Silva, and M. Monane. 1998. "A Longitudinal Analysis of Nursing Home Outcomes." *Health Services Research* 33 (4, part 1): 835–65.

Terza, J., A. Basu, and P. Rathouz. Forthcoming. "Addressing Endogeneity in Nonlinear Parametric Models: A Guide for Empirical Research in Health Economics." *Journal of Health Economics.*

Weech-Maldonado, R., L. Meret-Hanke, M. C. Neff, and V. Mor. 2004. "Nurse Staffing Patterns and Quality of Care in Nursing Homes." *Health Care Management Review* 29 (2): 107–16.

White, C. 2005. "Medicare's Prospective Payment System for Skilled Nursing Facilities: Effects on Staffing and Quality of Care." *Inquiry* 42 (4): 351–66.

Zhang, X., and D. C. Grabowski. 2004. "Nursing Home Staffing and Quality under the Nursing Home Reform Act." *Gerontologist* 44 (1): 13–23.

Supplementary Material

The following supplementary material for this article is available online:

Appendix A. 2SLS Stage-1 Results from Facility-Level Fixed Effects Regression: Predicting RN staffing and Skill Mix.

Appendix B. Sensitivity Analysis of Exclusion Restriction.

Table B. Sensitivity Analysis of Exclusion Criterion in Residual Inclusion Conditional Logit.

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1475-6773.2007.00803.x (this link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.