

© Health Research and Educational Trust
DOI: 10.1111/j.1475-6773.2006.00578.x

Addressing Measurement Error Bias in Nurse Staffing Research

David W. Harless and Barbara A. Mark

Objective. To assess the extent of measurement error bias due to methods used to allocate nursing staff to the acute care inpatient setting and to recommend estimation methods designed to overcome this bias.

Data Sources/Study Setting. Secondary data obtained from the California Office of Statewide Health Planning and Development (OSHPD) and the Centers for Medicare and Medicaid Services' Healthcare Cost Report Information System for 279 general acute care hospitals from 1996 to 2001.

Study Design. California OSHPD provides detailed nurse staffing data for acute care inpatients. We estimate the measurement error and the resulting bias from applying different staffing allocation methods. Estimates of the measurement errors also allow insights into the best choices for alternate estimation strategies.

Principal Findings. The bias induced by the adjusted patient days method (and its modification) is smaller than for other methods, but the bias is still substantial: in the benchmark simple regression model, the estimated coefficient for staffing level on quality of care is expected to be one-third smaller than its true value (and the bias is larger in a multiple regression model). Instrumental variable estimation, using one staffing allocation measure as an instrument for another, addresses this bias, but only particular choices of staffing allocation measures and instruments are suitable.

Conclusions. Staffing allocation methods induce substantial attenuation bias, but there are easily implemented estimation methods that overcome this bias.

Key Words. Nurse staffing, research methodologies, measurement error

The past decade has witnessed the publication of a number of studies examining the relationship between hospital registered nurse (RN) staffing and quality of care (American Nurses Association 1997, 2000; Kovner and Gergen 1998; Lichtig, Knauf, and Milholland 1999; Aiken et al. 2002; Kovner et al. 2002; Needleman et al. 2002; Cho et al. 2003; Unruh 2003; Mark et al. 2004). While the effects of nurse staffing on quality of care (measured as mortality, length of stay, and/or a variety of complications) are not entirely consistent, the studies begin to characterize the impact of nurse staffing on the various

outcomes. Caution is warranted, however, as all of these studies are based on nurse staffing estimates derived from datasets that are deficient in several respects. To overcome these deficiencies, researchers must rely on various ad hoc rules, particularly as they apply to the allocation of nurse staffing to the inpatient hospital setting. These methods are subject to errors in measurement, meaning that, at best, estimates of the effect of RN staffing on quality of care will be attenuated (biased toward 0). At worst, allocated staffing may serve as a proxy variable for actual staffing, obscuring the estimate of the parameter relating staffing level, and quality of care.

The state of California, however, through its Office of Statewide Health Planning and Development (OSHPD) requires hospitals to report productive staffing hours for specific service units (inpatient daily hospital services, ancillary services, and ambulatory services) so that inpatient acute care staffing may be calculated directly. Because of the level of detail provided in the California data, several authors have developed staffing allocation methods that rely on information from this data. The approach and purpose of this paper is different: we use the OSHPD inpatient staffing information to assess the performance of different staffing allocation methods and the size of the attenuation bias introduced by measurement error. We then summarize estimation strategies designed to overcome the attenuation bias, and provide evidence on whether the assumptions underlying these strategies are satisfied in the OSHPD data.

THE STAFFING ALLOCATION METHODS

Studies of the relationship between nurse staffing and quality of care require that data on nurse staffing be matched to the patients receiving care. For example, the number of nurses who work in long-term care or in ambulatory clinics should not be included in counts of RNs where the quality of care for hospitalized inpatients is being evaluated. Similarly, nurses who provide ancillary services should also be excluded. What is ultimately desired is an accurate count of the number of nurses who are providing care for patients

Address correspondence to David W. Harless, Ph.D., Department of Economics, Virginia Commonwealth University, Box 844000, Richmond, VA 23284. Barbara A. Mark, R.N., Ph.D., F.A.A.N, is with the School of Nursing, The University of North Carolina at Chapel Hill, Chapel Hill, NC.

receiving “daily hospital services.”¹ Although previous studies did not distinguish nurse staffing for inpatient acute hospital services from nurse staffing for inpatient services at ancillary cost centers (e.g., Kovner and Gergen 1998; Kovner et al. 2002; Needleman et al. 2002; Mark et al. 2004), our objective is to examine staffing levels in the inpatient acute areas so that a measure of staffing such as RN full-time employees (FTEs) per 1,000 inpatient days reflects an appropriate measure of staffing intensity.

There are two reasons for this approach. First, including RN FTEs involved in ancillary services combines unlike kinds of staffing: staffing for ancillary services should be measured with different metrics than staffing for hospital services. Second, even in the detailed California OSHPD data, it is not possible to distinguish inpatient and outpatient staffing at ancillary cost centers; all that is reported is productive hours at the ancillary service cost center. If staffing in inpatient acute cost centers is to be aggregated with staffing at ancillary service cost centers, there must be some auxiliary assumption (of the kind we wish to evaluate) about the allocation of this staff.

Calculation of recorded staffing level: Given the focus on inpatient acute cost centers, we begin by excluding staff employed in long-term care because staffing patterns differ substantially from those in the acute care facility.² In the California OSHPD data it is straightforward to exclude staff at long-term care cost centers as the productive hour information is reported directly. For hospitals in other states, one might begin with staffing data from the Centers for Medicare and Medicaid Services Provider of Services (POS) file, the American Hospital Association (AHA), or the financial and statistical reporting systems that exist in some other states. Among these sources, however, only the POS file distinguishes nurses employed in daily hospital services from those employed in skilled nursing/long-term care. Hence, the POS file, alone or in concert with other sources, would be used to obtain the number of RNs employed in the hospital.

The POS file, AHA, and most state reporting systems report the number of RN FTEs, which reflects both hours worked (productive hours) and paid time off (unproductive hours). The level of inpatient staffing can be defined based on the number of RN FTEs per 1,000 inpatient days (e.g., Mark et al. 2004) as we do in this paper, or by converting the FTEs to hours and defining inpatient staffing as the total number of hours per inpatient day (e.g., Needleman et al. 2002). An alternative is to estimate the number of productive hours and define inpatient staffing as the number of hours worked per inpatient day (American Nurses Association 2000). The California OSHPD data are unusual in that productive hours are reported, but not total hours. So that

our analysis provides sensible guidance for researchers using sources other than the California OSHDP data, we take productive hours and estimate FTEs.

In Table 1 we use an example California hospital to illustrate how we calculate the recorded staffing level, RN FTEs per 1,000 inpatient days, as well as the allocated levels under the different staffing allocation methods. For this hospital, the number of productive hours for direct payroll RNs is 281,464 and the number of productive hours for registry nursing personnel (which we

Table 1: Illustrating the Staffing Allocating Methods

Example for a California hospital

Direct payroll RN productive hours at hospital: 281,464
 Registry nurse productive hours: 53,924
 Direct payroll RN productive hours at inpatient acute cost centers: 175,714
 Registry nurse productive hours at inpatient cost centers: 47,074
 Inpatient days: 31,140
 Hospital RN FTEs: $(53,924 + 281,464 / 0.875) / 2,080 = 180.6$
 RN FTEs in inpatient acute cost centers: $(47,074 + 175,714 / 0.875) / 2,080 = 119.2$
 RN FTEs per 1,000 IPD: $119.2 / (31,140 / 1,000) = 3.83$

Administrative hours method

Total nursing administration hours: 1,421,781
 Nursing administration hours at inpatient acute cost centers: 891,404
 Share of total administration hours in inpatient acute cost centers: 0.627
 RN FTEs allocated to acute inpatient cost centers: $0.627 \times 180.6 = 113.2$
 RN FTEs per 1,000 IPD: $113.2 / (31,140 / 1,000) = 3.64$

Adjusted patient days method

Total outpatient revenue: \$38,032,017
 Total inpatient revenue (inclusive of ancillary services): \$106,172,177
 Adjusted patient days: $31,140 \times (1 + (38,032,017 / 106,172,177)) = 42,295$
 RN FTEs per 1,000 APD: $180.6 / (42,295 / 1,000) = 4.27$

Or

Total patient revenue: \$144,204,194
 Share of total inpatient revenue in gross patient revenues: 0.736
 RN FTEs allocated to inpatients: $0.736 \times 180.6 = 132.9$
 RN FTEs per 1,000 IPD: $132.9 / (31,140 / 1,000) = 4.27$

Revenue proportion method

Total patient revenue: \$144,204,194
 Patient revenue in acute inpatient cost centers: \$35,143,147
 Share of total patient revenue in inpatient acute cost centers: 0.244
 RN FTEs allocated to acute inpatient cost centers: $0.244 \times 180.6 = 44.1$
 RN FTEs per 1,000 IPD: $44.1 / (31,140 / 1,000) = 1.42$

Modified adjusted patient days method

RN FTEs per 1,000 IPD: $4.27 + (-0.19) \times (1 - 0.244) \times 4.27 = 3.66$

Modified revenue proportion method

RN FTEs per 1,000 IPD: $1.42 + 2.38 \times (1 - 0.244) \times 1.42 = 3.97$

RN, registered nurse; FTE, full-time employee; IPD, in-patient days; APD, adjusted patient days.

assume were all RNs) is 53,924. To convert to FTEs we assume that all registry nursing personnel hours were productive hours, a 40-hour work week ($40 \times 52 = 2,080$ paid hours in a year), and that 87.5 percent of total hours for direct payroll RNs were productive hours.³ For the hospital in Table 1, this implies there are 180.6 RN FTEs, of which 119.2 are at inpatient acute cost centers. Given 31,140 inpatient days at inpatient acute cost centers, the recorded staffing level is 3.83 RN FTEs per 1,000 inpatient days.

With data sources other than OSHPD, the starting point would be the number of RN FTEs employed at the hospital (180.6 for the example hospital in Table 1), and researchers must apply an allocation method to assign RNs to the provision of daily hospital services, excluding those who are employed in ancillary services and ambulatory services. There are several methods that have been applied. We describe each method, and, in Table 1, illustrate using data for the example hospital.

Administrative Hours Method. One allocation strategy uses Medicare Cost Report data on nursing administration hours (American Nurses Association 2000). Hospitals are required to report nursing administration hours by cost center on Worksheet B, Part I as part of the indirect cost step-down accounting to allocate overhead costs. In the administrative hours method, hospital RN FTEs are allocated to relevant cost centers in the same proportion used to allocate nursing administration hours to the cost centers. For the example hospital in Table 1, a total of 1,421,781 hours are reported for the nursing administration cost center across all hospital cost centers (excluding long term care cost centers). Of this total, 891,404 or 62.7 percent are reported in inpatient acute cost centers so 62.7 percent of the RN FTEs (113.2) are allocated to inpatient acute cost centers. Hence, the administrative hours method predicts 3.64 RN FTEs per 1,000 inpatient days.

Adjusted Patient Days (APD) Method. The most commonly applied (Kovner and Gergen 1998; Kovner et al. 2002; Mark et al. 2004) method for measuring overall staffing level is the APD method. The notion of APD assumes a common staffing “level” across hospital inpatient and outpatient cost centers given a particular normalization between inpatients and outpatients. The standard measure of volume for hospital inpatients is the inpatient day: the number of days that inpatients (excluding newborns in the nursery) are hospitalized. The APD concept assumes that outpatient visits can be normalized to an equivalent volume measure using the ratio of gross

outpatient and inpatient revenue: $APD = \text{inpatient days} \times [1 + (\text{outpatient revenue}/\text{inpatient revenue})]$. The measure of outpatient volume in inpatient-equivalent units is assumed to be inpatient days times the ratio of outpatient revenue to inpatient revenue. A hospital-wide measure of staffing level follows immediately: RN FTEs per adjusted patient day. Table 1 provides an example. The ratio of outpatient to inpatient revenue is $\$38,032,017/\$106,172,177 = 0.358$, so the total volume of outpatient care is assumed to be equivalent to $0.358 \times 31,140 = 11,155$ inpatient days. Hence, there are 4.27 RN FTEs per 1,000 APD. Although it may not be immediately apparent, note that this method allocates RN FTEs between inpatients and outpatients based on the proportion of gross patient revenues (inpatient revenue + outpatient revenue). In the example in Table 1, the share of inpatient revenue in gross patient revenues is 0.736 so 132.9 RN FTEs are allocated to inpatients resulting in an inpatient acute staffing level of 4.27 RN FTEs per 1,000 inpatient days. Needleman et al. (2001) note that the APD allocation method assumes equality of nurse staffing in inpatient and outpatient settings *per dollar of charges*.

Revenue Proportion Method. As the example above makes clear, the logic of APD method relies on the proportion of revenue and, by this logic, should be allocating RN FTEs to inpatient care *including* ancillary services. Our objective, however, was to evaluate methods of allocating RN FTEs to inpatient acute cost centers *excluding* ancillary cost centers. Applying the same logic based on proportion of revenue, we consider a variation on the method—allocating RN FTEs to inpatient acute care by the share of revenues at inpatient acute cost centers in gross patient revenue. For the example hospital, 24.4 percent of gross patient revenues are attributable to inpatient acute cost centers, so the method would allocate 24.4 percent of the RN FTEs (44.1) to these cost centers. The method predicts 1.42 RN FTEs per 1,000 inpatient days.

Modifications of the APD and Revenue Proportion Methods. Using the California OSHPD data, Needleman et al. (2001, 2002) examined RN staffing level inclusive of ancillary services assuming that ancillary cost center RNs were allocated to inpatients and outpatients according to their proportions of gross patient revenues. They found that the APD approach underestimated their measure of inpatient staffing and that the error was larger the greater the outpatient volume. They modified the APD method using a regression model

to estimate a correction factor. Applying their approach to the inpatient acute cost center staffing allocation problem we study, we estimate the parameter α in the equation

$$\text{RN Staffing} = \text{RN Staffing}_{\text{APD}} + \alpha(\text{RN Staffing}_{\text{APD}} \times (1 - \text{Inpatient Share})) + e \tag{1}$$

where RN Staffing is the recorded RN FTEs per 1,000 inpatient days at inpatient acute cost centers, $\text{RN Staffing}_{\text{APD}}$ is the allocated staffing level under the APD method, and Inpatient Share is the share of acute inpatient cost center revenue in gross patient revenue (0.244 for the hospital in Table 1). Given an estimate of the parameter α , the adjustment to the APD staffing level is larger the smaller the share of inpatient acute cost center revenue in gross patient revenue. Of course, a similar parameter can be estimated to modify the prediction of the revenue proportion method.

Table 2 contains estimates of α for five California OSHPD reporting cycles for the modifications to the APD and revenue proportion methods. Recall that α must be estimated using California data and then applied to other states (as in Needleman et al. 2002). Because we wanted to assess the performance of the different allocation methods as they would be applied, we did not use the parameter estimate from the same cycle to predict staffing level. Instead, to assess out-of-sample performance of the method, we use the α estimate from the previous reporting cycle. The data for the hospital in Table 1 are from reporting cycle 27; hence, we apply the α estimate from reporting cycle 26 (though the estimates for the two periods differ by only

Table 2: Estimates of α for the Modified Adjusted Patient Days Method and Modified Revenue Proportion Method*

<i>CA Report Cycle</i>	<i>Modified Adjusted Patient Days Method</i>	<i>Modified Revenue Proportion Method</i>
23	- 0.23 (0.01)	2.35 (0.05)
24	- 0.22 (0.01)	2.35 (0.05)
25	- 0.19 (0.01)	2.44 (0.06)
26	- 0.19 (0.01)	2.38 (0.06)
27	- 0.19 (0.01)	2.29 (0.06)

*Heteroskedasticity-robust standard errors (in parentheses) beneath each coefficient.

0.002). For the modified APD method, the estimate from reporting cycle 26 is -0.19 , and we modify the prediction of the APD method of 4.27 by subtracting $0.19 \times (1 - 0.244) \times 4.27 = 0.61$ for a predicted staffing level of 3.66 . The modification of revenue proportion method raises the prediction from 1.42 to 3.97 .

A SIMPLE MODEL TO ILLUSTRATE POTENTIAL BIAS

As it is clear from the examples above, each of the staffing allocation methods involves error, and the size and nature of the error can lead to substantial bias in estimating the effect of staffing on quality of care. Consider a simple model to illustrate the consequence of this measurement error in estimating the effect of nurse staffing on quality of care (Bound et al. 1994; Wooldridge 2002). Consider a simple regression model of quality of care, Y , and acute inpatient RN staffing level, X^* :

$$Y_i = \beta_0 + \beta_1 X_i^* + \varepsilon_i. \quad (2)$$

We assume that X^* is uncorrelated with ε . Although simplistic, this model with a single regressor provides benchmark estimates of bias.⁴ Our objective is to estimate β_1 , but we do not observe the true staffing level (except in the OSHPD data). Instead we observe

$$X_i = \delta_0 + \delta_1 X_i^* + u_i \quad (3)$$

where X_i is the staffing level calculated using one of the allocation methods. We assume that u_i is uncorrelated with ε_i . Substituting for X_i^* from (3) into equation (2) yields

$$Y_i = \left(\beta_0 - \beta_1 \frac{\delta_0}{\delta_1}\right) + \frac{\beta_1}{\delta_1} X_i + \left(\varepsilon_i - \frac{\beta_1}{\delta_1} u_i\right). \quad (4)$$

Equation (4) makes clear that the ordinary least squares (OLS) estimate of Y on X suffers from two possible sources of bias. First, if $\delta_1 \neq 1$ then allocated staffing level may be a *proxy* for the staffing level rather than a noisy measurement of the staffing level; hence, ignoring the measurement error, we estimate β_1/δ_1 , not β_1 . Second, even if $\delta_1 = 1$, the OLS estimate of β_1 will suffer from attenuation bias if X_i is correlated with the measurement error u_i .⁵ The size of the proportional bias toward 0 equals the coefficient γ_1 in the regression

$$(X_i - X_i^*) = \gamma_0 + \gamma_1 X_i. \quad (5)$$

Assuming $\delta_1 = 1$, the expected value of the OLS estimate is $\beta_1(1 - \gamma_1)$. For example, if $\gamma_1 = 1/3$, then the OLS estimate is expected to be two-thirds the size of β_1 . Finally, the classical measurement error model, in which u_i is assumed to be uncorrelated with X_i^* , represents a special case where

$$\gamma_1 = \frac{\sigma_u^2}{\sigma_{X^*}^2 + \sigma_u^2}. \quad (6)$$

That is, the proportional downward bias equals the ratio of the variance of the measurement error to the sum of the variances of the measurement error and the true staffing level.

DESCRIPTION OF SAMPLE

Our sample comprises general acute care hospitals in California from 1996 to 2001. Data on nurse staffing come from two sources. The first is the California OSHPD Hospital Annual Financial Data, which provides RN productive hours by cost center. This database contains desk-audited data collected from all acute care hospitals licensed by the State of California. We use data from the California OSHPD financial data files from reporting cycles 23–27, where, for example, report period ending dates for reporting cycle 23 are between June 30, 1997 and June 29, 1998. From the OSHPD files we obtain information on revenue, number of inpatient days for inpatient acute cost centers, and RN productive hours for hospital inpatient acute cost centers, ancillary services, and ambulatory services. The second source is Centers for Medicare and Medicaid Services' Healthcare Cost Report Information System (HCRIS) Worksheet B-1 files for HCRIS fiscal years 1996–2002 from which we obtain nursing administration hours at hospital inpatient acute cost centers, LTC/SNF cost centers, and total nursing administration hours.

We excluded the few instances in which hospitals had different report period beginning and ending dates in HCRIS and California OSHPD. Kaiser Foundation hospitals could not be included because they did not provide the patient revenue information necessary to apply the APD and proportion of revenue allocation methods. Following Needleman et al. (2002), we excluded the 251 instances in which hospitals had an average daily census of acute inpatients less than 20 in a reporting period. An additional eight potential observations could not be included because of missing data for patient revenue. Nursing administration hours were missing in 13 instances. Hospital RN productive hours were missing in a further five instances. Finally, 14

observations were excluded because the recorded hospital RN staffing information was obviously wrong. Though only the OSHPD data are detailed enough to indicate inpatient acute staffing, there are still instances where the data are, in our judgment, obviously incorrect. In our analysis, we compare staffing levels under the allocation methods to the recorded levels, assuming they are correct. But this comparison is nonsensical if the recorded values are obviously incorrect. We were conservative in these exclusions, however, deleting only those observations where hospital RN FTEs (adjusting for the number of days covered in a reporting period) rose by more than 40 percent or fell by more than 30 percent in a period, then returned to a level at or near the original level in the subsequent period (with no commensurate change in inpatient days).

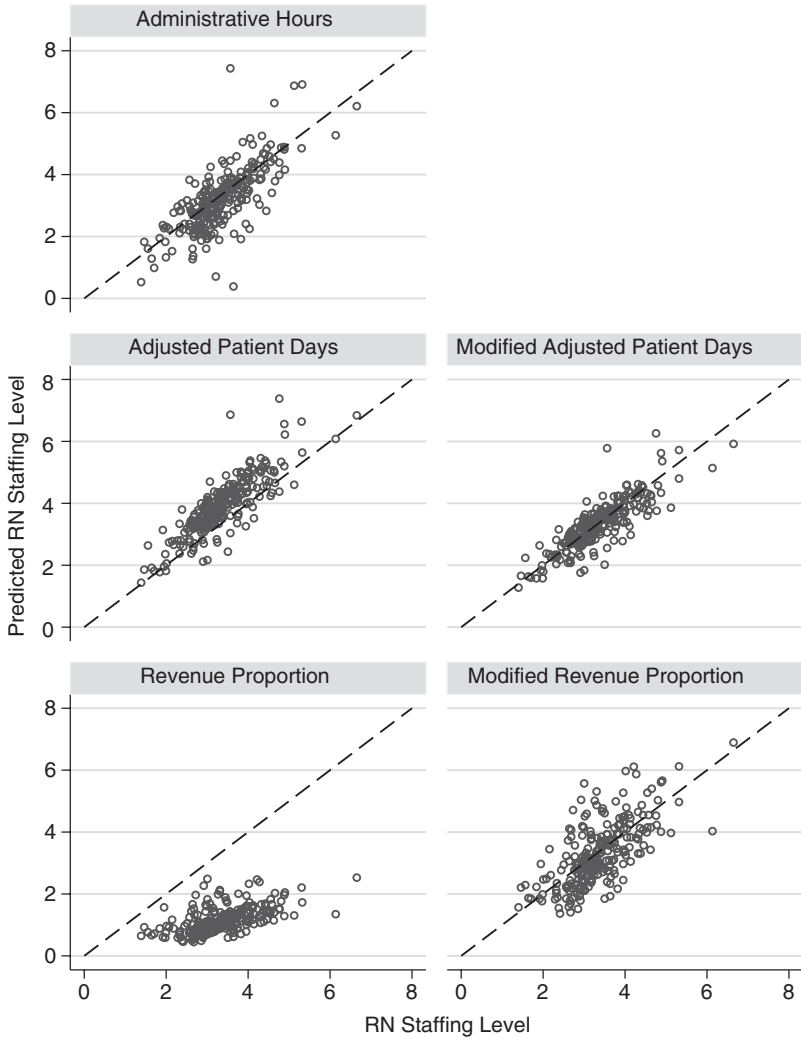
In practice, researchers might exclude observations if the allocated staffing level was judged to be an outlier. Part of the utility of the different allocation methods is the extent to which they can be applied without ad hoc deletions or imputations. As we wish to judge the performance of the methods avoiding subtleties in choices about whether a data point should be excluded, we did not exclude observations based on the allocated staffing levels. We did, however, exclude 11 observations because 100 percent of administrative hours were assigned to inpatient acute cost centers even though other sources of patient revenue were 78 percent or more of total patient revenues. (A mark against the administrative-hours method, as exclusions were required where the other methods could be applied.) A total of 279 hospitals are included in at least one reporting period.

RESULTS

We consider the performance of the allocation methods were they to be applied in a cross-sectional sample, which we take to be one California OSHPD reporting cycle. Further, so that we may gauge the stability and consistency of their performance, we report results for the four California reporting cycles 24–27.

Figure 1 plots allocated RN staffing level for each of the five methods versus recorded RN staffing level in reporting cycle 27. (Plots for other reporting cycles are similar.) Points falling on the 45° line indicate perfect agreement between recorded and allocated staffing levels. Figure 1 suggests that allocated staffing under the modified APD method is generally closer to the recorded levels, but there is substantial error under all the methods. As one way to illustrate the size of the errors, consider that the median of the absolute

Figure 1: Predicted Acute Inpatient RN Staffing Level for California Reporting Cycle 27, by Staffing Allocation Method.



Note: Acute inpatient RN staffing level is RN FTEs in inpatient acute cost centers per 1,000 acute inpatient days. Forty-five degree line indicates where predicted RN staffing level equals reported RN staffing level.

value of the error *as a percent of the recorded value of RN staffing* is 10.4 for the administrative hours method and 7.5 for the modified APD method; that is, a typical prediction is expected to differ from the recorded value by 10.4 or 7.5 percent.

Table 3 provides information on the difference between allocated and recorded staffing levels and an indication of the extent of the bias given the simple model described in the previous section. The first three columns of Table 3 indicate the reporting cycle, the number of hospitals included, and the mean and (in parentheses) the standard deviation of staffing level. The remaining columns give, for each staffing allocation method, the mean and standard deviation of the difference between allocated and recorded staffing levels (that is, $X_i - X_i^*$), and the estimate of δ_1 (from equation (3))⁶ and γ_1 (from equation (5)).

Consider first the estimates of δ_1 . The revenue proportion method produces a measure of staffing that must be characterized as a proxy variable for staffing rather than simply an error prone measure of staffing as the estimates of δ_1 are 0.28 – 0.30. All the other methods, however, have estimates of δ_1 that are quite close to 1. The modified APD method and modified adjusted revenue proportion method have δ_1 estimates that are sufficiently precise so that the parameter estimates can be judged to be statistically significantly different from 1 at standard significance levels.⁷ Nevertheless, in practical terms for the problem we address, the difference between 0.97 or 0.96 and 1.0 is small enough that we think it reasonable to maintain δ_1 approximately equal to 1 for all methods except the revenue proportion method.

Staffing levels from the administrative hours method and the APD method correspond to the classical measurement error model as the measurement error u_i is uncorrelated with the recorded value of RN staffing. Recall from the discussion above that when the measurement error is uncorrelated with the true value of RN staffing, the proportional bias toward 0 in estimating β_1 is given by either the slope coefficient γ_1 (equation (5)) or the ratio of the variance of the measurement error to the sum of the variances of the measurement error and the recorded staffing level (equation (6)). For example, in reporting cycle 24 for the administrative hours method, the slope coefficient measuring proportional downward bias is 0.39 which is equal to the bias calculated from the implied ratio of variances ($0.66^2 / (0.66^2 + 0.82^2) = 0.39$). The resulting bias in OLS estimates of the effect of staffing on quality of care in a model such as equation (4) is large: estimates of β_1 that are 39–45 percent too small using the administrative hours method and 30–32 percent too small using the APD method. We also note that the bias is expected to be still larger

Table 3: Performance of the Staffing Allocation Methods and Estimates of Bias

CA Report Cycle	N	RN FTEs Per 1,000 IPD		Administrative Hours Method		Adjusted Patient Days Method		Revenue Proportion Method		Modified Adjusted Patient Days Method		Modified Revenue Proportion Method		
		X-X*	δ_t	γ_t	X-X*	δ_t	γ_t	X-X*	δ_t	γ_t	X-X*	δ_t	γ_t	
24	236	3.33 (0.82)	-0.01 (0.66)	0.94 (0.39)	0.66 (0.54)	1.04 (0.31)	-2.19 (0.65)	0.30***	-0.34	0.97***	0.15	-0.10 (0.70)	0.96*	0.41
25	234	3.37 (0.77)	-0.12 (0.64)	0.93 (0.41)	0.56 (0.51)	1.06 (0.32)	-2.26 (0.63)	0.28***	-0.16	0.96***	0.17	-0.20 (0.73)	0.94***	0.47
26	232	3.34 (0.77)	-0.12 (0.65)	0.94 (0.41)	0.53 (0.50)	1.09 (0.32)	-2.20 (0.64)	0.28***	-0.12	0.98	0.20	-0.05 (0.73)	0.98	0.47
27	228	3.36 (0.79)	-0.14 (0.71)	0.98 (0.45)	0.55 (0.51)	1.02 (0.30)	-2.19 (0.65)	0.30***	-0.03	0.99	0.18	-0.05 (0.78)	0.98	0.49

*Significantly different from 1.0 at the .05 level.

**Significantly different from 1.0 at the .01 level.

***Significantly different from 1.0 at the .001 level.

RN, registered nurse; FTE, full-time employee.

in a multiple regression model. Under the classical measurement error model, Bound, Brown, and Mathiowetz (2001) gave a formula relating the bias to the R^2 statistic in a regression of true staffing on all other regressors in a multiple regression model. If this R^2 statistic was 0.50, then estimates of the staffing level parameter in a model for quality of care are 56–62 percent too small using the administrative hours method and 46–47 percent too small using the APD method.

The revenue proportion method, its modification, and the modified APD method all have a significant negative correlation between $X_i - X_i^*$ and recorded staffing levels, affecting the size (and even direction) of the measurement error bias. The most extreme case is for the revenue proportion method where the slope coefficient from a regression of $X_i - X_i^*$ on X_i is negative (e.g., -0.16 in reporting cycle 25) meaning that OLS results in an estimate of β_1/δ_1 that is too large. For the modified APD method and the modified revenue proportion method the negative correlation reduces the magnitude of the proportional bias toward 0. For example, in reporting cycle 24 the modified APD method has a ratio of variances of $0.45^2/(0.45^2+0.82^2) = 0.23$, but accounting for the correlation between $X_i - X_i^*$ and X_i^* , the proportional bias toward 0 is smaller, 0.15.

We conclude that the potential bias induced by measurement error is substantial and researchers should select estimation strategies that address this bias. There are several options.⁸ First, one might treat the California OSHPD data as *validation data* (Carroll, Ruppert, and Stefanski 1995; Bound, Brown, and Mathiowetz 2001). That is, one might use the California OSHPD recorded staffing levels to estimate a model to obtain predicted staffing levels conditional on the allocated staffing level *as well as* other regressors present in the quality of care equation which are assumed to be measured without error. Of course, in developing the modified APD method, Needleman et al. (2001, 2002) do something analogous in estimating an equation similar to equation (1). But this approach falls short of what is prescribed by statistical theory in several respects: The model for predicted RN staffing level should also include other regressors measured without error, be constructed such that the predicted values for staffing level are uncorrelated with the resulting measurement error, and when predicted staffing is substituted for actual staffing in the quality of care equation, standard errors should be adjusted appropriately (Carroll, Ruppert, and Stefanski 1995). As with the modified APD method and modified revenue proportion method, the validation data approach requires the assumption that the conditional distribution of staffing be the same for other states as it is for California. An additional concern is that the California

data, though from the best large database on staffing available to researchers, are not without error. Hence, some caution is warranted in its use as validation data.

Second, one might apply instrumental variable estimation, using one allocated staffing measure as an instrument for another allocated staffing measure included as a regressor in the quality of care equation. Again, using the simple model described in equations (2)–(4) as a benchmark, instrumental variable estimation of equation (4) produces a consistent estimate of β_1 if (i) the staffing measure included as a regressor has $\delta_1 = 1$, (ii) the measurement error, $X_i - X_i^*$, in this staffing measure is uncorrelated with the recorded staffing level X_i^* (mentioned above in the discussion of the measures of bias), and (iii) the measurement error in this staffing measure is uncorrelated with the measurement error in the instrumental variable (Bound, Brown, and Mathiowetz 2001). The revenue proportion method is not eligible to be included as a regressor as $\delta_1 < 1$, but this does not disqualify it as a potential instrument for other allocated staffing measures. Consistent with the differences in the bias measures mentioned earlier, panel A of Table 4 shows directly that the measurement error under the modified APD and modified revenue proportion methods, as well as the revenue proportion method, have statistically significant correlations with the recorded RN staffing level. The measures of staffing from the administrative hours method and the APD method appear to satisfy the assumptions to be the regressor included in (4), but panel B of Table 4 indicates that care should be exercised in choosing the instrumental variable. Whether the administrative hours method or the APD staffing measure is the regressor, only the revenue proportion method (or its modification in the case of the administrative hours method) satisfies the assumption that the measurement errors of the regressor and the instrument are uncorrelated. A high correlation in measurement errors for the adjusted patients days method and its modification is expected, but, somewhat surprisingly, measurement errors for the APD method and the administrative hours method have statistically significant positive correlations, at least in reporting cycles 24 and 26.^{9,10}

Finally, we note that the error term for equation (4), $(\varepsilon_i - \frac{\beta_1}{\delta_1} u_i)$, contains both ε_i , the error term in the original quality of care equation (2), and u_i , the measurement error. Condition (iii) for instrumental variable estimation restricts the measurement error in the instrumental variable from being correlated with u_i , the measurement error in the allocated staffing level included as a regressor. Table 4, panel B suggests that the allocated staffing under the revenue proportion method satisfies this condition. For this to be a valid instrument we emphasize that it (as well as the allocated staffing level included as a regressor) must also be uncorrelated with ε_i . We believe such an assumption is

Table 4: Tests of Conditions for Instrumental Variable Estimation

<i>Panel A. Correlation between Measurement Error and Recorded RN Staffing Level</i>					
<i>CA Report Cycle</i>	<i>Administrative Hours Method</i>	<i>Adjusted Patient Days Method</i>	<i>Revenue Proportion Method</i>	<i>Modified Adjusted Patient Days Method</i>	<i>Modified Revenue Proportion Method</i>
24	-0.08	0.06	-0.89***	-0.28***	-0.18*
25	-0.08	0.09	-0.87***	-0.23**	-0.19*
26	-0.07	0.14	-0.87***	-0.14	-0.16
27	-0.02	0.02	-0.84***	-0.23***	-0.12

<i>Panel B. Correlation of Measurement Error between Staffing Allocation Methods</i>					
<i>CA Report Cycle</i>	<i>Adjusted Patient Days Method</i>	<i>Measurement Error from</i>		<i>Modified Revenue Proportion Method</i>	
		<i>Revenue Proportion Method</i>	<i>Modified Adjusted Patient Days Method</i>		
<i>Measurement error from administrative hours method</i>					
24	0.24***	0.02	0.23**		-0.07
25	0.14	0.05	0.14		-0.02
26	0.17*	0.00	0.17*		-0.08
27	0.10	-0.02	0.08		-0.04
<i>Measurement error from adjusted patient days method</i>					
24		-0.00	0.93***		0.17*
25		-0.04	0.94***		0.09
26		-0.06	0.95***		0.14
27		0.07	0.96***		0.22**

*Significant at the .05 level.

**Significant at the .01 level.

***Significant at the .001 level.

RN, registered nurse

plausible as we know of no study that postulates that, say, the share of inpatient revenue in gross patient revenues determines quality of care.

CONCLUSION

Applied researchers, our noses deep in the data sources we rely on, are keenly aware of errors and inconsistencies we encounter. Though aware of the potential problems (e.g., attenuation bias) that can result from measurement error, we often plow on with conventional estimation strategies, explicitly, or implicitly, making convenient assumptions about the nature of the measurement error and hoping for the best. The staffing allocation methods we study

represent a different degree of measurement error because the available staffing data (outside of California OSHPD data) are at the hospital level and ad hoc rules are used to allocate staff to the inpatient setting. In this paper, we use data from California where inpatient acute care staffing is recorded directly and estimate the measurement error from applying various staffing allocation methods. We find that the measurement error in allocated staffing levels is large enough to induce significant bias. For example, in our benchmark simple regression model the APD method has measurement error large enough to cause the expected coefficient estimate for the effect of nurse staffing on quality of care to be 30–32 percent too small (and this attenuation bias becomes worse in a multiple regression model). Fortunately, there are easy to implement alternate estimation methods that can be applied to overcome this bias. Instrumental variable estimation may be applied, using the revenue proportion staffing measure as an instrument for staffing measured by APD or administrative hours (but some caution is warranted in using administrative hours as an instrument for APD or vice versa). Alternatively, the California OSHPD data may be treated as validation data to obtain predicted staffing levels which may be used as the measure of staffing without the resulting estimates suffering from attenuation bias.

ACKNOWLEDGMENTS

This research was supported by grant number R01HS10153 from the Agency for Healthcare Research and Quality. Thanks to Jack Needleman for his helpful comments.

NOTES

1. Specifically, we include the inpatient acute cost centers medical/surgical intensive care, coronary care, pediatric intensive care, neonatal intensive care, psychiatric intensive (isolation) care, burn care, other intensive care, definitive observation, medical/surgical acute, pediatric acute, psychiatric acute for both adult and adolescent and child, obstetrics acute, other acute care, nursery acute, and other daily hospital services.
2. Similarly, when we apply allocation methods using patient revenues or administrative hours we exclude revenue or hours attributed to long term care cost centers.
3. American Nurses Association (2000, p. 12) assumes 1,800 productive hours per FTE per year. Assuming 2,080 paid hours, this implies that 86.5 percent of total

hours were productive hours. Our estimate, 87.5 percent, differs slightly. The California OSHPD database does not separately report total hours for RNs but does report total hours for all hospital employees. In our sample the mean proportion of productive hours for hospital employees, weighted by total hours, was 87.5 percent, and this is the value that we apply.

4. Our benchmark model takes RN staffing levels to be exogenous, as does most of the existing research in this area (Kovner and Gergen 1998; Lichtig, Knauf, and Milholland 1999; Kovner et al. 2002; Needleman et al. 2002; Cho et al. 2003; Unruh 2003). In a longitudinal study, Mark et al. (2004) make the weaker assumption that staffing levels are “predetermined.” These assumptions have yet to be explicitly tested.
5. OLS produces unbiased estimates of β_1 when u_i is uncorrelated with X_i (correlated only with X_i^*), but this case does not occur in our sample for any of the allocation methods.
6. Recall that the staffing levels under the modified adjusted patient days method and the modified revenue proportion method are calculated using the recorded staffing level, and so should not differ significantly in their mean from the recorded staffing level. This implies that, for these methods, δ_1 should be estimated under the restriction that $\delta_0 = 0$.
7. For each method, four hypothesis tests are undertaken; hence, significance levels in Tables 3 and 4 are adjusted accordingly (Šidák 1967).
8. In addition to the validation data and instrumental variable estimation approaches we mention here, there is also the structural modeling or parametric approach (e.g., Fuller 1987; Hsiao 1989) where identification of model parameters depends on strong assumptions about the distribution of the measurement error.
9. In the circumstance of positive correlation in the measurement errors, instrumental variable estimation does result in tighter bounds on the parameter estimate: even though attenuation bias still exists, the attenuation bias is ameliorated compared with the OLS estimate of β_1 (Bound, Brown, and Mathiowetz 2001).
10. We undertook a similar analysis comparing performance of the staffing allocation methods in a longitudinal study. The conclusions concerning bias are similar to those drawn from the cross-sectional samples in Table 3. The longitudinal setting potentially yields another set of instrumental variables: if the measurement error for a staffing measure is serially uncorrelated, then lags of the measure can serve as instruments for the first-differenced staffing measure (Wooldridge 2002). Unfortunately, all the allocation methods have serially correlated measurement errors, closing off this possibility.

REFERENCES

- Aiken, L., S. Clarke, D. Sloane, J. Sochalski, and J. Silber. 2002. “Hospital Nurse Staffing and Patient Mortality, Nurse Burnout, and Job Dissatisfaction.” *Journal of the American Medical Association* 288 (16): 1987–93.

- American Nurses Association. 1997. *Implementing Nursing's Report Card: A Study of RN Staffing, Length of Stay and Patient Outcomes*. Washington, DC: American Nurses Publishing.
- . 2000. *Nurse Staffing and Patient Outcomes in the In-Patient Hospital Setting*. Washington, DC: American Nurses Association.
- Bound, J., C. Brown, G. Duncan, and W. Rodgers. 1994. "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data." *Journal of Labor Economics* 12 (3): 345–68.
- Bound, J., C. Brown, and N. Mathiowetz. 2001. "Measurement Error in Survey Data." In *Handbook of Econometrics*, vol. 5, edited by J. Heckman and E. Leamer, pp. 3707–843. Amsterdam: Elsevier Science.
- Carroll, R., D. Ruppert, and L. Stefanski. 1995. *Measurement Error in Nonlinear Models*. London: Chapman & Hall.
- Cho, S., S. Ketefian, V. Barkauskas, and D. Smith. 2003. "The Effects of Nurse Staffing on Adverse Events, Morbidity, Mortality and Medical Costs." *Nursing Research* 52 (2): 71–9.
- Fuller, W. 1987. *Measurement Error Models*. New York: John Wiley & Sons.
- Hsiao, C. 1989. "Consistent Estimation for Some Nonlinear Errors-in-Variables Models." *Journal of Econometrics* 41 (1): 159–85.
- Kovner, C., and P. Gergen. 1998. "Nurse Staffing Levels and Adverse Event Following Surgery in U.S. Acute Care Hospitals." *Image* 30 (4): 315–21.
- Kovner, C., C. Jones C, C. Zhan, P. Gergen, and J. Basu. 2002. "Nurse Staffing and Post-Surgical Adverse Events: An Analysis of Administrative Data from a Sample of U.S. Hospitals, 1990–1996." *Health Services Research* 37 (3): 611–29.
- Lichtig, L., R. Knauf, and D. Milholland. 1999. "Some Impacts of Nursing on Acute Care Hospital Outcomes." *Journal of Nursing Administration* 29 (2): 25–32.
- Mark, B., D. Harless, M. Mc Cue, and Y. Xu. 2004. "A Longitudinal Examination of Hospital Registered Nurse Staffing and Quality of Care." *Health Services Research* 39 (2): 279–300.
- Needleman, J., P. Buerhaus, S. Mattke, M. Stewart, and K. Zelevinsky. 2001. *Nurse Staffing and Patient Outcomes in Hospitals*. Boston: Harvard School of Public Health.
- . 2002. "Nurse-Staffing Levels and the Quality of Care in Hospitals." *New England Journal of Medicine* 346 (22): 1715–22.
- Šidák, Z. 1967. "Rectangular Confidence Regions for the Means of Multivariate Normal Distributions." *Journal of the American Statistical Association* 62 (318): 626–33.
- Unruh, L. 2003. "Licensed Nurse Staffing and Adverse Events in Hospitals." *Medical Care* 41 (1): 142–52.
- Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.