Staffing Subsidies and the Quality of Care in Nursing Homes

Andrew D. Foster and Yong Suk Lee*

June 12, 2013

Abstract

Concerns about the quality of state-financed nursing home care has led to the wide-scale adoption by states of pass-through subsidies, in which Medicaid reimbursement rates are directly tied to staffing expenditure. We examine the effects of Medicaid pass-through on nursing home staffing and quality of care by adapting a two-step FGLS method that addresses clustering and state-level temporal autocorrelation. We find that pass-through increases staffing by 4.4% in higher quality nursing homes and improves quality of care in lower quality nursing homes by an amount equivalent to one sixth the inter-quartile range of the quality distribution.

Keywords: Medicaid, wage pass-through, nursing home, direct care staff, quality of care, FGLS

JEL Codes: H51, I18, J20, C23

---

* Foster: Brown University, Providence RI 02912 (email: afoster@brown.edu); Lee: Williams College, Williamstown MA 01267 (email:yong.s.lee@williams.edu). We thank Christian Hansen for providing code based on his 2007 Journal of Econometrics article. This work was supported by the National Institute on Aging grants R01-AG026465 and P01-AG027296. We thank Jonathon Skinner, Vincent Mor, Zhanlian Feng, Orna Intrator, Sylvia Kuo and other participants in the P01-AG027296 grant. We also thank seminar participants at Brown University's Applied Microeconomics Lunch Seminar for helpful comments. Tim Squires implemented the procedure as a full fledged Stata command.

I. INTRODUCTION

A difficult issue that arises in terms of the provision of public programs is how to keep down costs while at the same time ensuring that providers provide a level of quality that meets appropriate standards. User assessment of quality may not provide sufficient incentive if users do not have sufficient choice among alternative providers, if they cannot readily monitor quality or if providers are not otherwise able to benefit from improvements in quality in terms of higher prices (because of limits on out of pocket costs) or higher demand (if the reimbursement cost is close to the marginal cost of care). Direct monitoring of output quality is difficult and must often at least rely in part on self-reported data. In some cases it is more feasible to subsidize or regulate inputs that are thought to have a direct impact on quality. However, in such cases there is an obvious concern that providers may respond to such incentives by cutting back on other needed inputs, thus leading to little impact on quality.

A particularly clear example of this problem arises in the context of nursing home care. Medicaid helps the poor gain access to old age long-term care by directly reimbursing nursing homes. However, Medicaid reimbursement rates are considerably lower than private-pay rates and in many cases barely cover the cost of care provided. Hence, nursing homes with a high percentage of Medicaid patients may have little incentive to increase quality beyond a level that would trigger unwanted attention (Scanlon 2001). Such concern, along with the sense that labor input is a primary determinant of the quality of care in nursing homes, has led to the wide-scale adoption by states of direct subsidies of nursing staff in the form of pass-through policies, in which Medicaid reimbursement rates are directly tied to staffing expenditure. States that adopt pass-through policies earmark additional funds to the direct-care staff, the majority of whom are Certified Nurse Assistants (CNAs), and generally require nursing homes to use the pass-through to increase wages, benefits, or the number of staff.

In 1999, only 6 states (California, Delaware, Michigan, Minnesota, South Carolina, and Wisconsin) had a wage pass-through policy in place. In the five-year period that followed, 17 additional states introduced a wage pass-through. Figure 1 shows the distribution of CNA hours over the 6 year period for Florida, New York, and Massachusetts combined. These are the only three states that adopted the policy in 2002 and maintained it until the end of our sample period. They comprise some of the largest nursing home markets in the nation. A mean shift in CNA hours starting in 2002 is evident, suggesting a potential policy effect. Moreover, the distribution becomes bimodal, pointing to potential heterogeneous effects. The main question we ask in this paper is whether wage pass-through increases CNA staffing hours and ultimately the quality of care in nursing homes. An increase may occur at the mean but can also be more pronounced at certain parts of the distribution as suggested in Figure 1. In addition to identifying mean pass-through effects, understanding why heterogeneous effects can occur theoretically and identifying such effects empirically is an important goal of the paper.

A series of papers have examined the relationship between the Medicaid reimbursement rate and the quality of nursing home care. Gertler (1989, 1992) studies the effect of an increase in Medicaid payment rate on nursing home quality and access. He finds, utilizing 1980 New York data, that higher Medicaid payment rates increase Medicaid admissions but decrease the quality of care. In these papers, demand for nursing home care by Medicaid patients exceeds supply as nursing home capacity is constrained by the Certificate of Need (CON) program, and nursing homes provide equal service to both private pay and Medicaid patients. In such settings, a nursing home has the incentive to reduce cost by lowering service quality when it chooses to displace a private pay patient with a Medicaid patient. Similarly, Nyman (1988) tests the excess demand hypothesis using 1979 Wisconsin data and finds that higher reimbursement rates are associated with lower quality of care.

However, as Grabowski (2001) and Grabowski et al. (2004) points out, the nursing home market condition has changed considerably since the 1980s, with various long-term care options becoming available to patients. The excess demand assumption for nursing home markets has become less reasonable with national nursing home occupancy rate falling steadily to 84.8% by 2007 (Harrington et al., 2008). Using national data for 1996, Grabowski (2001) finds a positive effect of reimbursement on quality of care. Cohen and Spector (1996) also find a positive effect of reimbursement on staffing but no effect on quality. Our paper contributes to this literature by using national nursing home data for six consecutive years to examine the effect of a relatively new wage pass-through program as well as the overall Medicaid reimbursement rate on nursing home staffing and quality.

We believe this is the first study to develop and test a theoretical framework of nursing home responses to wage pass-through. Previously in Feng et al. (2010) we examined some results on staffing by applying a procedure outlined in this paper. Here we motivate and explain the underlying estimation procedures and show the efficiency gains of using different procedures in multi-level panel data. Furthermore, we present a theory of nursing home behavior that helps explain the policy's heterogeneous impact on staffing and quality. Baughman and Smith (2010) examine the impact of wage pass-through on the wages of direct care workers utilizing the Surveys of Income and Program Participation (SIPP) data and find that wages increase within states but are unable to find a wage increase when using individual fixed effects. We focus on nursing home level data and the differential response of different types of nursing homes to the policy, specifically on how these policies affect patient outcomes.

Other studies that have addressed staffing issues point to the importance of worker effort and labor substitutability in studying the efficacy of a wage pass-through program. Currie et al. (2002) find that hospital takeover decreases hospital competition on quality and thus allows hospitals to lower labor costs by increasing nurse work load, rather than decreasing wages. Cawley et al. (2004) study factor substitution in the nursing home industry and find that higher nursing home wages are associated with lower staffing and greater material usage, particularly, physical restraints. These studies suggest that the overall impact of a wage pass-through policy would not only depend on how nursing homes respond in terms of staffing but also on how the changes in nurse effort result from the change in wages. In order to assess both the direct and indirect effects of wage pass-through on quality we embed an efficiency wage model with staffing constraints into a standard model of nursing home behavior.[1] Our model indicates that lower quality nursing homes for which staffing constraints are binding increase wages when subjected to a wage pass-through. On the other hand, higher quality nursing homes that are not constrained legally in terms of staffing numbers, increase staffing with the wage pass-through. This heterogeneous response implies that efficiency per worker increases at the lower quality nursing homes but decreases at the higher end, ultimately resulting in a larger increase in quality at the lower end nursing homes.

Empirically, we utilize a newly constructed panel data set of nursing home characteristics and state policies from the 48 continental U.S. states over a six-year period from 1999 to 2004. Empirical analysis of policies is hampered by the fact that variation in policies occurs at the level of the state over time and outcomes may be influenced by aggregate state level shocks that are autocorrelated across time. In order to address the empirical difficulties pertaining to multi-level panel data we adapt a Two-step FGLS method proposed by Hansen (2007). We simplify

[1] Norton (2000) provides an overview of nursing home models introduced in the literature.

Hansen's procedure so that the two-step FGLS method can be conveniently implemented. Our approach, in particular, corrects for the bias associated with measuring the AR(1) coefficient in short panels, while also controlling for facility-level fixed effects. This method serves as an alternative to conventional Difference in Difference methods that cluster at the state-level to obtain robust estimates against the clustering and intertemporal autocorrelation problems (Bertrand et al. 2004). We show analytically and by simulation that our Two-step FGLS estimates reduce standard errors by about 25-40%. In our empirical analysis we observe a larger efficiency gain of over 50%. Two-step FGLS can be generalized to any policy analysis utilizing multi-level panel data and in such cases provides more efficient estimates than conventional Difference in Difference estimates.[2]

We perform our analysis on all urban Medicaid certified nursing homes as well as samples stratified based on each nursing home's average Medicaid share of beds. We also non-parametrically estimate the heterogeneous policy impact over the distribution of Medicaid share of beds. Consistent with the model, we find that pass-through increases staffing with the impact being stronger in high technology nursing homes, where staffing hours increase by 4.4% . In terms of quality, we find that pass-through improves quality by an amount equivalent to one sixth of the inter-quartile range of the quality distribution, especially at low technology nursing homes with a high Medicaid share of beds.

In the next section we develop a simple efficiency wage model with staffing constraints to examine wage pass-through effects. Section III describes our empirical strategy and the benefits of the two-step FGLS estimation procedure. Section IV describes the data as well as how we construct a measure of quality of care. Section V illustrates our empirical results and Section VI concludes.

II. A Model of Nursing Home Behaviour

Pass-through can affect quality of care through (1) increases in staffing or (2) changes in worker effort[3]. Staffing can change because the price of labor nursing homes face decreases with

[2] The authors along with Tim Squire have created a single line Stata command "xtfear" which implements the procedure outlined in the paper. The Stata ado file "xtfear" is available upon request to the authors.

[3] As Weiss (1990) points out quality of workers can improve because of the change in distribution of workers hired (adverse selection) or changes in individual effort (incentive effect or moral hazard). The model in this section is tailored to the incentive effects model because it concisely explains the heterogeneous effects without having to go through a more complicated job search model of turnover.

a wage pass-through. Worker effort can change because one's effort depends on the wage one receives from the nursing home as well as one's opportunity wage, as typically stipulated in an efficiency wage model. The overall impact of the policy depends on how nursing homes respond not only to the change in market wage but also to the changes in nurse effort resulting from the change in market wage. In order to understand both the direct and indirect effects of wage pass-through on quality and to model the effect of wages on worker productivity we embed an efficiency wage model with staffing constraints into the nursing home model generally used in the literature (Gertler 1989, 1992, Norton 2000).

We assume that nursing homes accept both private-pay and Medicaid patients, provide the same quality of care to both patient types, and do not discriminate admission by patient type. The long-term care (LTC) market today is increasingly competitive, with more and more non-nursing home LTC providers, such as assisted living and other community-based services, competing with traditional nursing homes. The national average nursing home occupancy rate has fallen below 85% in recent years.[4] This allows us to consider a nursing home's Medicaid patient load as exogenously determined by the local market condition. Nursing home markets are local, with counties being a reasonable approximation, given patterns of funding and resident origin (Foster and Rahman 2011; Grabowski 2008). Nursing homes can charge different prices to private pay patients depending on the quality of care each facility provides. Finally, nursing homes face staffing constraints as mandated by federal and state law. Staffing standards vary widely across states with certain states setting statutes in terms of the nurse to patient ratio and others in nurse hours per patient day. Also, statutes in some states concern direct-care staff in general without specifically restricting CNA staffing hours.[5] In our optimization problem, we generalize the staffing mandate as an inequality constraint on staffing.

An important element of our research is the distinction between hiring greater staffing and raising the quality of output per worker. It is unclear, for example, that an increase in staffing will improve patient well-being if the effect of this increase in staffing, given other conditions in the market, ends up reducing worker effort. To capture this idea we incorporate an efficiency wage model. In particular, the quality of care is produced in terms of efficiency labor units, i.e.,

[4] The national nursing home occupancy rate in 2001 was 85.9% and has been falling steadily to 84.8% by 2007 (Harrington et al., 2008). Gertler's (1989, 1992) studies are based on New York's nursing home market in the early 1980s when occupancies were well over 95%.

[5] Harrington (2008) describes nursing home staffing standards for all states in detail.

$q=e(w_1,w_2)l$, where q is quality, e is efficiency per worker, and l is staffing hour per patient. Efficiency is determined by the wage received from the nursing home, w_1, and the market wage, w_2. Efficiency increases as nurses earn higher wages ($e_{w1} > 0$) but decreases as the market wage increases holding own wage constant ($e_{w2} < 0$). A nursing home providing care of quality q and charging v per private-pay patient attains $F(v,q)$ number of private pay patients where we assume $F_q > 0$ and $F_v < 0$. For Medicaid patients, nursing homes earn the Medicaid reimbursement rate s, which is determined by the state, and the number of patients G_0, which we take as exogenously determined by the structure of the local market.[6] Nursing homes receive the additional wage pass-through amount of pw_1 where p is the subsidy rate on the nurse wage w_1. Nursing homes choose the private pay patient price v and the number of staff l and their wage w_1 to maximize profit[7]

$$\Pi = vF(v,q;\delta)+ sG_0 + pw_1l\ G_0 – w_1l(F(v,q;\delta)+ G_0\ ) \qquad (1)$$

subject to $l \geq l_{min}$ where l_{min} is the mandated minimum staffing per patient. δ is a technology parameter where higher δ enables nursing homes to attract more private pay patients ($F_\delta >0$), resulting in a lower share of Medicaid patients. We are primarily interested in the effects of introducing wage pass-through p on labor demand l and quality of care q.

We first heuristically discuss the impact of a wage pass-through introduction and then illustrate using a parametric example. For unconstrained nursing homes, that is, nursing homes of high technology that optimally choose labor above l_{min}, the first order conditions $\Pi_v = 0$, $\Pi_l = 0$ and $\Pi_{w1} = 0$ determine optimal private pay price, staffing, and wage. However, low technology nursing homes that would optimally choose staffing below l_{min} are mandated to choose l_{min} and determine a wage accordingly. Depending on whether a nursing home is constrained by the staffing mandate, the introduction of a wage pass-through will have different implications. The first order effect of a wage pass-through is to lower the price of labor for nursing homes resulting in an increase in demand for labor, and if nursing home labor is inelastically supplied to the relevant market, an increase in the market wage w_2. Nursing homes that are not constrained by the staffing mandate will increase staffing and adjust their own wages based on their first order conditions. However, nursing homes that were constrained will still choose l_{min}, unless the

[6] The number of Medicaid patients is, of course, potentially endogenous with respect to nursing home quality, but less so, we believe because of the nature of Medicaid financing. E.g., those with Medicaid funding have overall fewer options for care than private pay patients.
[7] Most nursing homes are for-profit. Moreover, for our purposes, and given a reasonable set of objectives for non-profit nursing homes the key behaviors of interest do not vary by for profit status.

subsidy was large enough to move them above l_{min}, and thus utilize the full pass-through amount to increase wages for the staff. The increase in own wages would be large for these constrained nursing homes, resulting in increased worker efficiency, and ultimately, quality of care. Though staffing increases in high technology nursing homes, worker efficiency may drop because the own wages may not rise as much as the market wage rises, resulting in a possible decrease in worker effort.

To graphically illustrate these twofold effects more clearly we consider the following parametric example. In particular, we assume

$$F(v,q) = \frac{\delta q}{v^{\alpha+1}} \qquad (2)$$

$$e(w_1, w_2) = \left(\frac{w_1 - \eta_1}{w_2 - \eta_2} \right)^{\nu} \qquad (3)$$

$$w_2 = 4l \qquad (4)$$

and for parameter values we choose η_1=1, η_2=1, α=0.9 and ν=0.9.[8] To gain intuition into the underlying process, we first hold the technology parameter fixed (δ = 40) and illustrate how the introduction of pass-through returns different results based on whether the staffing mandate is binding or not. This nursing home's optimal choice of wage and staffing is 10 and 1.08 without the pass-through policy. Figures 2A and 2B represent the nursing home behavior when it is unconstrained and constrained, respectively. If the staffing mandate is 1 then a given nursing home would choose its optimal wage and staffing level as in Figure 2A. However, as illustrated in Figure 2B, if the staffing mandate is higher, say at 1.2, then it would choose a lower wage of 7.8. The introduction of a pass-through of p=0.2 shifts the nursing home's labor demand curve out and the unconstrained nursing home chooses a wage of 10 and a higher staffing level of 1.2. The fact that wage remains at 10 is a consequence of the separability in equation (3), which we impose for expositional simplicity. However, as Figure 2B illustrates the constrained nursing home still has to choose staffing at the mandated level of 1.2, but is now able to raise its wage to 10. Thus, when wage pass-through is introduced (1) unconstrained nursing homes will increase staffing while constrained nursing homes remain at the mandate level and (2) constrained nursing homes will increase wages more than unconstrained nursing homes.

---

[8] Equation (4) is a simple way to characterize the market wage that captures aggregate demand effects.

Now we solve the model by allowing δ to vary and fixing l_{min} at 1. We assume δ to represent a specific local market of patients, whereas, the market for nurses is determined at the state level. This exercise illustrates how the various nursing homes of different technology levels across markets respond to wage pass-through. Figures 3A and 3B present nursing home staffing and wage decisions by technology level. As discussed before, staffing hours increase at high δ but remain unchanged at low δ. On the contrary, wages increase at low δ but remain unchanged at high δ. As a result, quality in lower technology nursing homes improves with pass-through but quality slightly decreases in higher technology nursing homes as illustrated in Figure 3C. The quality improvement is more pronounced in low technology nursing homes because it sees a considerable increase in staff efficiency due to the rise in own wages relative to that of the market. Quality change in high technology nursing homes is minimal because the increase in staffing is offset by the drop in staff efficiency. Staff efficiency may drop because, given the separability, own wages are fixed and market wages rise. Lastly, Figure 3D shows how the share of Medicaid patient decreases as the technology parameter increases. The Medicaid share is higher with wage pass-through in place but as higher technology nursing homes are able to capture more private pay patients, the Medicaid share in those nursing homes steadily declines with δ.

The parametric model illustrates how the impact of wage pass-through on the staffing level and quality of care differ depending on the technology parameter. In other words, if there are two local patient markets, one characterized by high and the other by low technology nursing homes, the low technology market will increase wages without hiring additional nursing staff but the high technology market will hire more nurses without increasing individual wages. In the empirical analysis that follows, we examine the average effect of wage pass-through on staffing and quality but also proxy technology with the Medicaid share of beds to examine heterogeneous policy effects.

III. Empirical Research Design

This section discusses the empirical research design that enables us to efficiently estimate state policy effects using facility-level panel data over relatively short time horizons by implementing a simplified version of Hansen's (2007) FGLS procedure. We first define the parameters of interest and describe our estimation strategies.

III.A Potential Problems for Estimation

Consider

$$y_{ist} = X_{ist}\beta + Z_{st}\gamma + \mu_s + \eta_t + \varepsilon_{st} + v_i + u_{ist} \qquad (5)$$

where y_{ist} is the outcome variable (e.g., staffing level or quality of care), X_{ist} is a vector of facility level variables including county characteristics, Z_{st} is a vector of state policy variables, μ_s represents state specific unobservables that determine staffing hour (e.g., other labor or health policies relevant to staffing level), η_t is a vector of year dummies, v_i reflects facility-level unobserved variables that determine our outcome (e.g., management style), ε_{st} and u_{ist} represent state-level and facility-level idiosyncratic factors. u_{ist} is assumed to be independent of all observable and unobservable factors in (5). We assume that ε_{st} is independent of Z_{st}, the state policy variables, and later in section V.A examine this assumption using a hazard model.

The parameter of interest is γ. However, there are several problems that complicate the consistent and efficient estimation of γ. First, facilities in the same state at a particular point in time are likely to face common shocks, such as, economic conditions or unobserved state policies. Thus, facilities in the same state and time cannot be treated as independent, i.e., $E(u_{ist}\, u_{jst})\neq 0$. This is the so called clustering problem. Second, these common shocks are likely to be serially correlated within states over time. So residuals at the state-level cannot be treated as independent over time, i.e., $E(\varepsilon_{st}\varepsilon_{st-k})\neq 0$. Lastly, there is an endogeneity problem because facility-level characteristics such as share of patients in Medicaid or patient acuity may be influenced by state-year economic or environmental shocks, i.e., $E(X_{ist}\varepsilon_{st})\neq 0$. Hence, regressions at the facility level may lead to biased estimates of these characteristics and thus lead to misleading policy effects.

A potential solution and widely used estimation method is to employ facility level Difference in Difference regression with Huber-White standard errors clustered at the state level (Bertrand et al. 2004). This method is robust to the clustering and autocorrelation problem but will return biased estimates when the state-time level shocks are correlated with facility characteristics. Even if we can assume $E(X_{ist}\varepsilon_{st})=0$, this method is potentially inefficient to the extent autocorrelation can be captured with a simple AR process and have low power against the two-step Feasible GLS method suggested by Hansen (2007). In this paper, we adapt and slightly modify Hansen's procedure so that the two-step FGLS method can be easily implemented. The main modification is in the estimation of the AR(1) parameter which we discuss below. This method can be generalized to any policy analysis susceptible to serially correlated state-year shocks and in such cases provides more efficient estimates than conventional Difference in Difference estimates.

III.B Estimation Strategy: Two-step OLS and Two-step FGLS

As Hansen (2007) illustrates we can group the variables that are constant at the state-time level and rewrite equation (5) as

$$y_{ist} = X_{ist}\beta + c_{st} + v_i + u_{ist} \qquad (6)$$

where

$$c_{st} = Z_{st}\gamma + \mu_s + \eta_t + \varepsilon_{st}. \qquad (7)$$

c_{st} captures the policy component of the outcome as well as the state and time level fixed effects. In order to cancel out the facility level fixed effects in (6) we demean at the facility level so that

$$\dot{y}_{ist} = \dot{X}_{ist}\,\beta + \dot{c}_{st} + \dot{u}_{ist} = \dot{X}_{ist}\,\beta + \textstyle\sum_s \sum_t \dot{c}_{st}\,\delta_{st} + \dot{u}_{ist} \qquad (8)$$

where a dot indicates that variables are demeaned at the facility level (e.g., $\dot{y}_{ist} = y_{ist} - \bar{y}_i$) and δ_{st} are state-time dummies. We can estimate $\dot{c}_{st}$ with state-time indicators as in (8). Once we estimate $\dot{c}_{st}$, we can estimate γ through equation (7) employing a Difference in Difference regression at the state level with Huber-White standard errors clustered by state. Clustering at the state level and not at the state-time level allows estimates to be robust to serial correlation in the state-time shocks. Note that $\dot{c}_{st} = c_{st} - \bar{c}_s$, so using $\dot{c}_{st}$ instead of c_{st} in equation (7) will not change the estimates of γ since $\bar{c}_s$ is subsumed in the state fixed effects. In other words, the actual estimation is

$$\dot{c}_{st} = Z_{st}\gamma + \tilde{\mu}_s + \eta_t + \varepsilon_{st} \qquad (9)$$

where $\tilde{\mu}_s = \mu_s - c_s$. We denote this estimation strategy Two-step Difference in Difference (DD). Two-step DD returns consistent estimates of gamma but is inefficient due to the serial correlation in the residuals. The Two-step FGLS procedure, as documented in Hansen (2007), imposes an AR(p) structure on the residuals and further improves upon the Two-step DD procedure in terms of power. In this paper we model the serial correlation in ε_{st} with an AR(1) process and perform a standard FGLS estimation by first estimating the AR(1) coefficient ρ and then performing OLS on the ρ-differenced variables. We believe the serial correlation in the state-time level shocks in most policy or program analyses can be captured well with an AR(1) process even when there are further lags in the AR process. In section III.D we show that assuming AR(1) still has substantial gains in efficiency even when the underlying process is of higher order. Hence, the empirical strategy we pursue is to extract an AR(1) process in the state-time shocks and further cluster at the state level in the regressions to allow for any remaining correlation we do not capture with the AR(1) structure.

Estimating the AR(1) coefficient can be problematic. In an NxT panel data with short T, estimating the AR coefficient by regressing the residuals, i.e., regressing ε_{st} on its lags will return biased estimates. The next section provides the intuition and an illustration of this bias. Hansen's (2007) main contribution is in obtaining unbiased estimates of the AR coefficients in panels with short T. Our methodological contribution is in implementing a quick and intuitive method to obtain the AR(1) coefficient that does not require iteration as Hansen's approach normally would.

For the Two-step FGLS estimation, we demean variables at the state level in (9) to eliminate the state fixed effects $\tilde{\mu}_s$ so that

$$\dot{c}_{st} - \bar{\dot{c}}_s = (c_{st} - \bar{c}_s) - (\bar{c}_s - \bar{c}_s) = (Z_{st} - \bar{Z}_s)\gamma + (\eta_t - \eta) + (\varepsilon_{st} - \bar{\varepsilon}_s) .$$

Note that (1) demeaning $\dot{c}_{st}$ simply returns $\dot{c}_{st}$, (2) η is simply subsumed in the constant term in the actual regression, and (3) $\bar{\varepsilon}_s = 0$ from the assumption of a stationary AR(1) process of state-time shocks $\varepsilon_{st} = \rho\, \varepsilon_{st-1} + e_{st}$ where e_{st} is White noise. Hence we get the following state-year level equation:

$$\dot{c}_{st} = \dot{Z}_{st}\gamma + \eta_t + \varepsilon_{st} . \quad (10)$$

With an unbiased estimate of ρ, we then perform the following ρ-differenced regression to retrieve the Two-step FGLS estimates of γ.

$$\dot{c}_{st} - \rho\, \dot{c}_{st-1} = (\dot{Z}_{st} - \rho\, \dot{Z}_{st-1})\, \gamma + \xi_t + e_{st} \quad (11)$$

where $\xi_t = \eta_t - \rho\eta_{t-1}$. Even after we net out the AR(1) process there could be remaining correlation within state across time, i.e., it may still be that $E(e_{st}e_{st-k}) \neq 0$. Hence, we still cluster standard errors by state. Also, in order to correctly include the first observation of each state in equation (11), the first observation of each state is normalized by $\sqrt{1-\rho^2}$ so that the variance of ε_{s1} equals the variance of e_{st}.

We have outlined the Two-step procedures (both DD and FGLS) used in our analysis. Lastly, we point out the advantage of estimating a multilevel panel (facility level with state policy) in two steps. The multilevel structure creates potential endogeneity between time varying state shocks and facility level variables i.e., $E(X_{ist}\varepsilon_{st})\neq 0$. Many DD analyses assume away this possibility and simply drop ε_{st} in the estimation of equation (5). However, there are many instances that we should suspect endogeneity. In our case, we have reasons to believe that year varying state economic shocks (e.g., unemployment) or weather shocks affect the share of Medicaid patients or patient acuity in nursing homes. The two step process allows us to separate these two effects and allows us to work around this endogeneity.

III.C Unbiased Estimation of the AR(1) Coefficient

In this section we show how one can compute an unbiased estimate of the AR(1) coefficient. First we outline the source of bias in estimating the AR(1) coefficient in short panels. Consider our second stage equation (10) $\dot{c}_{st} = \dot{Z}_s\gamma + \eta_t + \varepsilon_{st}$ and the AR(1) process $\varepsilon_{st} = \rho\varepsilon_{st-1} + e_{st}$. Suppose we regress the estimated residual $\hat{\varepsilon}_{st}$ on its lag $\hat{\varepsilon}_{st-1}$ to estimate ρ. Effectively we are regressing $\varepsilon_{st} - \frac{1}{n}\sum_{t=1}^{T}\varepsilon_{st}$ on its lag. If we had a long panel with large T the term $\bar{\varepsilon}_s = \frac{1}{n}\sum_{t=1}^{T}\varepsilon_{st}$ would converge to zero. However, in short panels this term becomes non negligible so that $\operatorname{cov}(\varepsilon_{st},\varepsilon_{st-1}) \neq \operatorname{cov}(\hat{\varepsilon}_{st},\hat{\varepsilon}_{st-1})$ and ultimately $\rho \neq \hat{\rho}$. A simple example with T=2 illustrates this more concretely. In the case of T=2, $\hat{\varepsilon}_{s1} = \frac{(\varepsilon_{s1} - \varepsilon_{s2})}{2}$ and $\hat{\varepsilon}_{s2} = \frac{(\varepsilon_{s2} - \varepsilon_{s1})}{2}$. So the correlation in the estimated residuals is -1 regardless of actual value of ρ. In general the bias $\hat{\rho} - \rho$ is negative with the absolute value of the bias decreasing in T.

Hansen (2007) develops a general method that calculates the magnitude of the bias and subtracts away the bias to get an unbiased coefficient estimate of general AR(p) processes. In this paper, we develop a simple method of moments approach that estimates the AR(1) coefficient. Hansen's method has the advantage of being more generalizable to p^{th}-order processes, but our focus on AR(1) enables us to derive a simple one-step formula for the autocorrelation coefficient. A drawback of our method is that we use differences in the estimation and hence lose one degree of freedom for each state. However, comparison of our estimate with Hansen's estimate when T=6 show almost no difference.

Our estimation method utilizes (a) the fact that the stationarity of an AR(1) process implies that the differenced process,

$$\Delta\varepsilon_{it} = \rho\,\Delta\varepsilon_{it-1} + \Delta e_{it}, \quad (12)$$

is also stationary, and (b) the Yule-Walker equations. Denoting the r^{th} autocovariance function $\mathrm{cov}(\Delta\varepsilon_{t+r}, \Delta\varepsilon_t) = h_r$ and $\mathrm{var}(e_{it}) = \sigma_e^2$ we get the following two Yule-Walker equations:

$$(12) \times \Delta\varepsilon_{it-1} \longrightarrow h_1 = \rho h_0 - \sigma_e^2$$

$$(12) \times \Delta\varepsilon_{it} \longrightarrow h_0 = \rho h_1 + (2\text{-}\rho)\, \sigma_e^2$$

Solving for ρ we get

$$\rho = 1 + 2\frac{h_1}{h_0}.$$

Hence, the sample counterpart of the AR(1) coefficient is

$$\hat{\rho} = 1 + 2\frac{\mathrm{cov}(\hat{\varepsilon}_{t+2} - \hat{\varepsilon}_{t+1}, \hat{\varepsilon}_{t+1} - \hat{\varepsilon}_t)}{\mathrm{var}(\hat{\varepsilon}_{t+1} - \hat{\varepsilon}_t)}.$$

We use this $\hat{\rho}$ in equation (11) to complete the two-step FGLS estimation. As mentioned before we employ either Huber-White standard errors clustered by state or jack-knife standard errors at the state level in order to ensure robustness against potential clustering problems that may still exist after accounting for the AR(1) serial correlation.

III.D Efficiency Gains of Two-step FGLS

Before we take the above estimation strategies to our data we first examine the efficiency gains from the Two-step FGLS both analytically and by simulation. First, consider $y_{it} = x_{it}\beta + u_{it}$ where T=2 and the residuals are serially correlated by an AR(1) process. If the autocorrelation for the residual is ρ and the autocorrelation for the regressor is κ then the percentage reduction in variance of FGLS relative to OLS as shown in the Appendix is

$$1 - \frac{1-\rho^2}{1-\kappa^2\rho^2}.$$

For estimated parameters from the data with ρ=0.7 and κ=0.5 this amounts to a 42% reduction in variance or 24% reduction in standard errors of the estimated β. Note that doubling the number of states should reduce variance by 50% or standard errors by 29%. Therefore, implementing the Two-step FGLS results in similar efficiency gains as having a panel with twice as many states.

Next, we perform a Monte Carlo simulation to examine the efficiency gains from Two-step FGLS relative to Two-step DD and standard DD estimation. The data generating process randomly generates facility variable x, policy variable z, and parameters roughly consistent with the data for 50 states, over 6 or 12 time periods, with 0-100 facilities per state. We set the true β, the coefficient on z, equal to one and carry out the exercise 100 times implementing (1) DD estimation with standard errors clustered at the state level, (2) Two-step DD with standard errors clustered by states, (3) Two-step FGLS with standard errors clustered by states, and (4) Two-step FGLS with standard errors jack-knifed at the state level. The results in Table 1 clearly demonstrate the efficiency gains of the Two-step FGLS approach. In the case of a panel with 6 time periods, the Two-step FGLS results in standard errors that are about 40% smaller than the general DD estimation. The efficiency gain increases with a longer panel of T=12 and remains similar even when we estimate an AR(2) error structure with an AR(1) process. Moreover, the fact that the standard deviations of the estimated coefficients are similar to the average standard errors of the coefficients illustrates that the estimation procedure provides correct standard errors.

Both the analytic exercise and the simulations show that Two-step FGLS has considerable efficiency gains relative to conventional DD methods, resulting in 30-45% reduction in standard errors depending on the true underlying AR structure. These results are consistent with Hansen's finding that the FGLS confidence interval decreases by 44% compared to the OLS case based on a simulation over resampled CPS-MORG data.

IV. The Data and Constructing a Measure of Quality

IV.A Data

The main data for this study is facility-year level panel based on the Centers for Medicare and Medicaid Services' (CMS) Online Survey Certification and Reporting (OSCAR) system from 1999 through 2004. OSCAR contains annual self reported facility-level information, including staffing, organizational characteristics and aggregate resident conditions, for all Medicare/Medicaid certified nursing homes in the U.S. If surveys were conducted more than once a year for a given facility, we use the one closest to year end. We exclude hospital-based nursing homes, which primarily serve short-stay, post-acute Medicare patients. Staff in these nursing homes is often provided by the affiliated hospital. We also exclude Medicare only certified facilities which do not serve Medicaid patients and rural nursing homes because staffing patterns and the structure of long term care market differ considerably between urban and rural areas. (Feng et al. 2008). Finally, we exclude the small number of facilities located outside the 48 contiguous U.S. states. State policy and local nursing home market information are then merged into the facility level data. State policy, such as Medicaid wage pass-through payments, average Medicaid per diem rates, and the use of case-mix reimbursement, were collected from surveys of state Medicaid offices conducted by Brown University's Community Health Department. Lastly, the Bureau of Health Professions' county-level Area Resource File provides local nursing home market characteristics.

Table 2 reports the variables used in the analysis and the descriptive statistics over the study period. The unit of analysis is the nursing home facility. Staffing level is defined as the total average hours per resident day (HPRD) by Certified Nursing Assistants (CNAs). We focus on CNAs, because they are the primary target of the wage pass-through program. Furthermore, there is little concern about substitution of CNAs with other, more expensive types of direct care staff, such as RNs and LPNs. Our other outcome variable, quality, requires more subtle construction. We discuss in detail how we define a measure for quality in the next section.

There are three state policy variables in our analysis. One of the key variables is the wage pass-through policy to nursing homes by Medicaid. Each state differs in how wage pass-through is actually implemented. Some states set the compensation in dollar amounts per staff hour/per patient day. Other states set compensations rates or allocate a pool of money for the purpose of wage pass-through. For the empirical analysis we focus on the effect of introducing the pass-through policy and employ a single dummy variable identifying the provision of wage pass-through payments in a given state and year over the period 1999-2004.[9] Another policy variable we are interested in is the annual state-average Medicaid reimbursement rate (per bed day) to nursing homes, a variable that the previous literature has focused on. The reimbursement rate not only varies among states but also varies within state over year depending on annual budgets. We use inflation adjusted 2004 dollar amounts in the analysis. Lastly, we include an indicator variable for the presence of a case-mix adjusted reimbursement policy. States with a case-mix policy adjust Medicaid reimbursement rate based on the average acuity of patients in the nursing home.

As in previous studies, we use county variables to proxy for nursing home market variables. We use the average number of empty beds per nursing home in the county as a

[9] Pass-through was introduced in the following states by year as follows: California, Delaware, Michigan, Minnesota, South Carolina, and Wisconsin in 1999; Kansas, Maine, Montana, Texas, Virginia, and Vermont in 2000; North Dakota, Rhode Island, and Wyoming in 2001; Arizona, Florida, Massachusetts, Maryland, and New York in 2002; Louisiana and South Dakota in 2003; and Georgia in 2004.

measure of excess supply and of market competition. We include the number of Registered Nurses (RNs) and Licensed Practical Nurses (LPNs) per hospital bed, number of nursing home beds per thousand population over 65, the Medicare managed care penetration rate to capture the variations in local demand and supply factors pertaining to nursing home workforce, and a standardized hospital wage index to account for regional differences in the purchasing power of Medicaid payments and the price of medical and nursing services.

To control for facility level characteristics we include an acuity index which combines resident's activity of daily living (ADL) dependencies and special treatment measures. Also included are whether the facility employed or contracted for a nursing practitioner or physician assistant, and whether the facility operated an Alzheimer unit. In addition, we control for the payer mix of residents in each facility, as indicated by the percent of Medicaid residents.

IV.B Measuring Quality of Care

Input based measures, such as labor or capital usage, have often been used to measure nursing home quality of care but this method assumes a restrictive technology. Regulatory violations have also been used to proxy quality, but such measures may capture different regulatory environments or market conditions. In this paper, we use direct output measures. We develop a quality measure utilizing patient health outcomes that are most responsive to labor input. Four measures of patient outcome, the activities of daily living (ADL) decline rate, the restraint rate, the pressure ulcer (PU) worsening rate, and the persistent pain (PP) rate were constructed from the Minimum Data Set (MDS), a federally mandated resident assessment data, as outlined in Mor et al. (2011).[10] Not all measures may well represent nursing home quality of care, since some conditions may worsen due to the patient's inherent condition rather than the care given by nurses. The pressure ulcer worsening rate is considered as a good measure of care in the literature because bedsores are generally preventable and treatable conditions (Grabowski 2008). Our approach in deciding which outcome variables to use is to examine which variables best represent quality as an outcome of labor input.

Suppose A_{it} and B_{it} are two different measures of patient outcome and can be expressed as

[10] ADL measures how well residents perform normal daily activities, such as, bathing, dressing, eating, walking, using the toilet, etc. Restraint rate measures usage of belts, vests, pelvic ties, specialized chairs or bed side rails in nursing homes to prevent wandering, Pressure ulcer worsening rate measures the proportion of residents with bedsores and persistent pain rate is the ratio of patients with worsening or persistent pain problems.

$$A_{it} = \alpha_i + (L_i + R_{it}) + \varepsilon_{it}$$

$$B_{it} = \beta_i + \eta(L_i + R_{it}) + v_{it}$$

where α_i and β_i are facility i's inherent capacity to produce outcomes A and B, and η is a relative productivity parameter. L_i is some staffing level that facility i maintains consistently and R_{it} is variable staffing susceptible to economic conditions or state policies. The first differences return

$$A_{it+1} - A_{it} = (R_{it+1} - R_{it}) + (\varepsilon_{it+1} - \varepsilon_{it})$$

$$B_{it+1} - B_{it} = \eta(R_{it+1} - R_{it}) + (v_{it+1} - v_{it}).$$

This implies that the change in patient outcome overtime is captured by the change in stochastic staffing as well as some idiosyncratic shock. So if both A and B are good measures of quality the differences, $A_{it+1} - A_{it}$ and $B_{it+1} - B_{it}$, would be correlated strongly. Table 3 provides the correlation among the first differenced outcome variables. The ADL decline rate and the PU worsening rate have a positive correlation of 0.13 which is significantly larger than all other combinations of variables. Hence, these two variables appear to be the better measures of quality produced by staffing input. Rather than using one patient outcome to proxy quality of care we can reduce the variance of the idiosyncratic shocks by taking a weighted sum of these two variables. In order to determine the appropriate weights, we examine the loading factors in a factor analysis involving (i) only ADL decline rate and PU worsening rate and (ii) all four patient outcome variables. As Table 4 indicates, when we include only the two variables we get one common factor with identical loadings. Also, when we include all four variables we get two common factors and the loading factors on ADL decline rate and PU worsening rate are nearly identical as well. Hence, we construct a quality index with identical weights so that

$$q = 0.5(Z_{ADL} + Z_{PU})$$

where q is the quality of care measure, Z_{ADL} is the standardized ADL decline rate, and Z_{PU} is the standardized PU worsening rate. Table 5 illustrates the distribution of the quality index by states. A more negative value implies higher quality. The state with the highest average level of quality is North Dakota at -0.35, the lowest is Indiana at 0.35, and the inter-quartile difference is 0.22. This number will help us interpret the coefficients later in the empirical results.

V. EMPIRICAL RESULTS

V.A Hazard Analysis of Pass-through Policy Implementation

Before we examine the effect of policy on our outcome variables we perform a simple hazard analysis to examine whether the implementation of pass-through policy by states may be considered exogenous, i.e., whether $E(Z_{st}\varepsilon_{st})=0$. Though we can not directly test for this exclusion restriction, we can examine whether policy enactment is related to observables in our data. We estimate whether initial staffing level or quality of care affects the probability that a certain state enacts a pass-through policy in the following years. In particular, we estimate a logit regression on the following semi-parametric equation, which allows flexibility in the hazard function:

$$y_{st} = \mathrm{X}_{s1999}\beta + d_{2000} + d_{2001} + d_{2002} + d_{2003} + d_{2004} + \varepsilon_{st} \,,$$

where y_{st} is a binary variable indicating whether state s had the policy in year t, X_{s1999} is a vector of relevant covariates including staffing level and quality level aggregated to the state level for base line year 1999, and d's are dummy variables equal to one in the subscript year. Table 6 presents the results. Column I includes 1999 average CNA staffing and quality, the other two state policy variables, and year indicators. Neither staffing hours nor quality is significant. We then include the facility and market level variables in columns II through IV. The coefficients on all variables as reported on Table 6 are not significant even at the 10% level, suggesting that initial levels of staffing and quality in 1999 did not determine the subsequent decision of pass-through policy take up. Though not reported in the table, the coefficient on the facility and market level variables are also not significant. These results suggest that policy endogeneity is not to be a problem in our following analyses.

V.B Empirical Findings on Staffing and Quality

Panel A of table 7 presents the effect of state policies on log CNA hours under various estimation models. Column I reports coefficients and standard errors from the conventional DD model with Huber-White standard errors clustered at the state level. We find that pass-through increases staffing level by about 2.5% but is not statistically significant. As mentioned earlier, this estimate may be biased due to unobservable state level shocks and is potentially inefficient, because the estimation does not utilize the policy autocorrelation in the residuals over time within states. Columns II-IV report the second stage results of Two-step DD (equation (9)) and

Two-step FGLS (equation (11)) where the first stage is facility level OLS with state-time level fixed effects (equation (8)). The component of staffing that varies at the state-time level net of facility level characteristics is carried on to the second stage. Column III uses Huber-White standard errors clustered by states and column IV estimates the standard errors by jack-knifing over states. The standard error on pass-through decreases substantially to 0.013 from 0.024 when we implement two-step DD. The point estimate drops a little to about 0.019 but the impact is still not statistically significant. However, when we employ two-step FGLS the standard error on pass-through drops even further to about 0.009 and the point estimate becomes statistically significant at the 10% level. The Huber-White standard errors and jack-knifed standard errors return nearly identical results. The estimated AR(1) coefficient is 0.79 indicating that there is strong autocorrelation in the error term. By incorporating this autocorrelation the two-step FGLS results in over 50% reduction in standard errors compared to conventional DD estimates. The efficiency gain from actual data is even better than that of the Monte Carlo simulations in section III.C. Based on the results in columns III and IV, we find that wage pass-through increases CNA staffing level by about 1.8%. Also, results indicate that Medicaid reimbursement rate and case mix policy on average do not have any significant effect on staffing level.

Panel B of Table 7 reports results on quality as defined in section IV.C. In interpreting the numbers we remind the reader that a lower number implies better quality. We employ the same DD and FGLS estimation procedures as before, except that Column 1 and the first stage for the two-step procedures are estimated at quarterly intervals since patient outcome measures are quarterly variables. Also, policy variables are lagged one year because quality variables are generated using relative changes from previous periods (Mor et al. 2011). We find that pass-through on average increases quality significantly at a magnitude equivalent to about one seventh of the interquartile distribution. Also, when we compare columns III and I, the reduction in standard errors on pass-through is once again over 50%, with an estimated AR(1) coefficient of 0.88. The change in signs on the pass-through coefficient and the fact that the Medicaid reimbursement rate effect goes a way with the two-step FGLS procedures suggests that the conventional DD estimates in column I may also be suffering from endogeneity problems.

Overall, results in Table 7 indicate that the wage pass-through policy, on average, increases both staffing level and quality of care. These magnitudes may seem weak from a policy point of view but as our theoretical model predicts the policy impact could vary across nursing homes. We next analyze heterogeneous effects by examining subsets of the data and performing non-parametric estimation.

V.C Heterogeneous Effects in Stratified Samples

In this section, we estimate the policy effect on samples stratified by nursing home average Medicaid share of beds. We proxy nursing home's inherent technology level with the Medicaid share of beds to capture the relation where higher technology nursing homes attract more private pay patients resulting in a lower share of Medicaid patients. We use the Medicaid share of beds, rather than patients, because total patient numbers and occupancy rates are susceptible to outside factors, whereas, the number of total beds is a reasonably fixed feature of the nursing home. We stratify the sample into quartiles and then estimate the policy effects on each quartile using the Two-step FGLS procedure. One concern is that Medicaid share of beds is distributed differently by state. If we stratify over the whole sample of nursing homes some states may have nursing homes predominantly in the first quartile and others in another quartile. In order to overcome this problem we stratify the sample into quartiles within each state.

Table 8 presents the results of the stratified analysis. We present standard DD estimates, Two-step DD estimates, and Two-step FGLS estimates. The 1st quartile represents the highest technology nursing homes with the lowest Medicaid share of beds and the 4th quartile represents nursing homes of the lowest technology with highest Medicaid share of beds. Panel A of Table 8 indicates that wage pass-through increases staffing levels by about 2.9% in the first quartile, with significance just over the 10% level, and increases staffing significantly by 4.4% in the second quartile. Note that this effect is much larger than the 1.7% found over the whole sample. The policy effect that is concentrated on the better quality nursing homes are dispersed over the whole sample in Table 7.

On the other hand, the impact of pass-through on quality of care is more pronounced in the 3rd and 4th quartile, the lower technology nursing homes with higher Medicaid share of beds. Panel B of Table 8 indicates that the introduction of pass-through improves the quality measure in the 4th quartile by 0.036 and in the 3rd quartile by 0.027. The estimates are less precise being significant at the 15% level. Note that this 0.036 is about one sixth of the inter-quartile difference in quality as documented in Table 5, which is similar to moving up the ranks by about 4 states. Also, a $10 increase in Medicaid reimbursement rate improves quality by about 0.028 and 0.025

in the 3rd and 4th quartile respectively. This increase in quality attributed to the increase in Medicaid reimbursement rate in lower technology nursing homes suggest, as is evident in the occupancy rate, that there may not be excess demand in this subset of nursing homes contrary to what was found in the literature that examines 1980 nursing homes.

These heterogeneous results are consistent with the predictions of the parametric model. We empirically find that the introduction of wage pass-through policy increases staffing without changing quality in the higher technology nursing homes but increases quality without changing staffing in the lower end.

V.D Non-parametric Estimation of Heterogeneous Effects

The above analysis on stratified samples provides only snap shots of the heterogeneous effects. In order to graphically illustrate how the policy effects evolve along the full range of Medicaid share of beds, we next perform non-parametric estimations. We weight the data using a kernel around a Medicaid share value and then implement the two-step FGLS procedure to retrieve one set of coefficients on pass-through and Medicaid reimbursement rate. We repeat this exercise over the distribution of Medicaid share and then plot the coefficients. Figures 4A and 4B illustrate the effect of pass-through on staffing and quality, respectively. Figures 4C and 4D present the impact of Medicaid reimbursement rate.

Consistent with what we see in Table 8, Figure 4A shows that the pass-through effect on staffing is strongest and significant at the higher technology nursing homes. The point estimate around the hump is about 0.3 which is between the 1st and 2nd quartile estimates in Table 8. Note that the standard error bands expand towards the edge, giving less precise estimates, even though point estimates remain large. Also consistent with Table 8, pass-through increases quality at the lower technology nursing homes in Figure 4B, with the largest magnitude being around 0.04. These figures confirm that the stratified estimation in Table 8 effectively represents the heterogeneous impacts well.

Figures 4C and 4D also point to interesting heterogeneous effects. The point estimates on the Medicaid reimbursement rate, a continuous measure, are more precise than those on pass-through and the heterogeneous effects are quite stark. Higher Medicaid reimbursement rate increases staffing only at nursing homes with low Medicaid share of beds, a result that is consistent with the fact that the model imposes a minimum staffing constraint that only binds for low quality homes. The zero effect is precisely measured for the rest of the nursing homes. Similarly, Medicaid reimbursement rate increases quality only at nursing homes with the highest Medicaid share of beds. Though our model in this paper focuses on the pass-through policy and neutralizes the Medicaid reimbursement rate effects, a similar story can be told by making Medicaid admission endogenous in the model. In such a model, an increase in the reimbursement rate would lead to more Medicaid admission in nursing homes with high Medicaid vacancy, most likely the high technology nursing homes. Staffing demand would increase in these nursing homes resulting in higher staffing. On the other hand, low technology nursing homes where Medicaid beds are at capacity will see little or no increase in Medicaid admission.[11] So staffing would not change and the increased reimbursement rate would result in an increase in nurse wages. The efficiency wage set up could then explain the heterogeneous quality effects depicted in Figure 4D.

VI. Conclusion

A central question in health economics and public economics more generally is whether and under what conditions it is possible to achieve desired outcomes by subsidizing inputs. The problems are both theoretical and empirical. From a theoretical perspective it is important to assess not only the extent to which the incentives of the provider are allied with those of policy makers and opportunities for substitution across different margins, but also to consider possible general equilibrium effects of these policies. From an empirical perspective there are limited sources of policy variation that cover a sufficiently large area to create general equilibrium responses but a sufficiently small area to provide reasonable temporal and spatial variation. The states in the US provide a reasonable source of policy variation, but the need to control for autocorrelated state-wide shocks substantially reduces power, at least when standard state-clustering methods are used.

In this paper we address both issues in the context of an analysis of the effects of Medicaid pass-through regulations on nursing home staffing and patient outcomes. From a theoretical perspective we develop a simple parametric model that captures variation in policy responses arising from whether or not existing staffing constraints are binding, incorporates endogenous effort, and allows for simple general equilibrium feedback. The model suggests that

[11] As before, this heuristics hinges on the assumption that nursing homes are not at full capacity.

staffing-constrained nursing homes will pay higher wages as a result of pass-through while unconstrained homes will increase staffing but not wages. Given overall wage increases, worker effort may decline, particularly in higher technology nursing homes.

From an econometric perspective we implement a simplified version of Hansen's model for small-T panels with autocorrelated shocks and fixed effects. In particular, we develop a simple consistent estimator of the AR(1) parameter and then show, using simulation, that this alternative method leads to substantial improvements in efficiency relative to the usual approach of simply clustering on states. Implementation of the procedure using nursing home data leads to reductions of standard errors of 50%, even when clustering is used in addition to Feasible GLS to ensure robustness against autocorrelation that is greater than AR(1).

Overall, the results support the nuanced predictions of the model. Wage pass-through led to greater staffing in the higher technology nursing homes but had little effect on staffing on lower technology nursing homes. On the other hand, wage pass-through led to improvements in the quality of care in the lowest technology nursing homes. These results indicate that pass-through can be an effective instrument for improving nursing home quality among nursing homes already constrained by existing staff constraints. More generally, the results point out the importance of input subsidies that work in tandem with existing constraints to limit possibilities of substitution.

References

Amemiya, Takeshi, "A Note on a Random Coefficient Model," *International Economic Review,* 19 (1978), 793-796.

Baughman, Reagan A. and Kristin Smith, "The Effect of Medicaid Wage Pass-Through Programs on the Wages of Direct Care Worker," *Medical Care,* 48 (2010), 426-432.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-In-Differences Estimates?" *Quarterly Journal of Economics,* 119 (2004), 249-275.

Currie, Janet, Mehdi Farsi, and W. Bentley Macleod, "Cut to the Bone? Hospital Takeovers and Nurse Employment Contracts," NBER Working Paper no. 9428 (2002).

Cawley, John, David C. Grabowski, and Richard A. Hirth, "Factor Substitution and Unobserved Factor Quality in Nursing Homes," NBER Working Paper no. 10465 (2004).

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

Feng, Zhanlian, David C. Grabowski, Orna Intrator, Jacqueline Zinn, and Vince Mor, "Medicaid Payment Rates, Case-Mix Reimbursement, and Nursing Home Staffing," *Medical Care,* 46(2008), 33-40.

Feng Zhanlian, Yong Suk Lee, Sylvia Kuo, Orna Intrator, Andrew D. Foster, and Vincent Mor, "Do Medicaid Wage Pass-Through Payments Increase Nursing Home Staffing?" *Health Services Research*, 45(2010), 728-747.

Foster, Andrew D., and Momotazur Rahman, "Racial Segregation and Quality of Care Disparity in US Nursing Homes" , manuscript, Brown University, 2011.

Gertler, Paul J., "Subsidies, Quality, and the Regulation of Nursing Homes," *Journal of Public Economics,* 38 (1989), 33-52.

——— "Medicaid and the Cost of Improving Access to Nursing Home Care," *Review of Economics and Statistics,* 74 (1992), 338-345.

Grabowski David C., "Does an Increase in Medicaid Reimbursement Rate Improve Nursing Home Quality?" *Journal of Gerontology,* 56 (2001), 84-93.

——— "The Market for Long-Term Care Services," *Inquiry,* 45 (2008), 58-74.

——— "Medicaid Reimbursement and the Quality of Nursing Home Care," *Journal of Health Economics*, 20 (2001), 549-569.

Grabowski, David C., Zhanlian Feng, Orna Intrator, and Vincent Mor, “Recent Trends in State Nursing Home Payment Policies,” *Health Affairs Web Exclusive*, W4 (2004), 363-73.

Harrington, Charlene, “Nursing Home Staffing Standards in State Statutes and Regulations,” *mimeo* (2008).

Harrington, Charlene, Helen Carillo, and Brandee W. Blank, “Nursing Facilities, Staffing Residents and Facility Deficiencies, 2001 through 2007,” *mimeo* (2008).

Hansen, Christian B., “Generalized Least Squares Inference in Panel and Multilevel Models with Serial Correlation and Fixed Effects,” *Journal of Econometrics*, 140 (2007), 670-694.

Mor, Vincent, Andrea Gruneir, Zhanlian Feng, David C. Grabowski, Orna Intrator, and Jacqueline Zinn, “The Effect of State Policies on Nursing Home Resident Outcomes,” *Journal of American Geriatrics Society*, 59 (2011), 3-9.

Norton, Edward C., “Long-term Care,” in A.J. Cuyler, J.B. Newhouse (Eds.), *Handbook of Health Economics, Vol.1*, (Amsterdam: Elsevier, 2000).

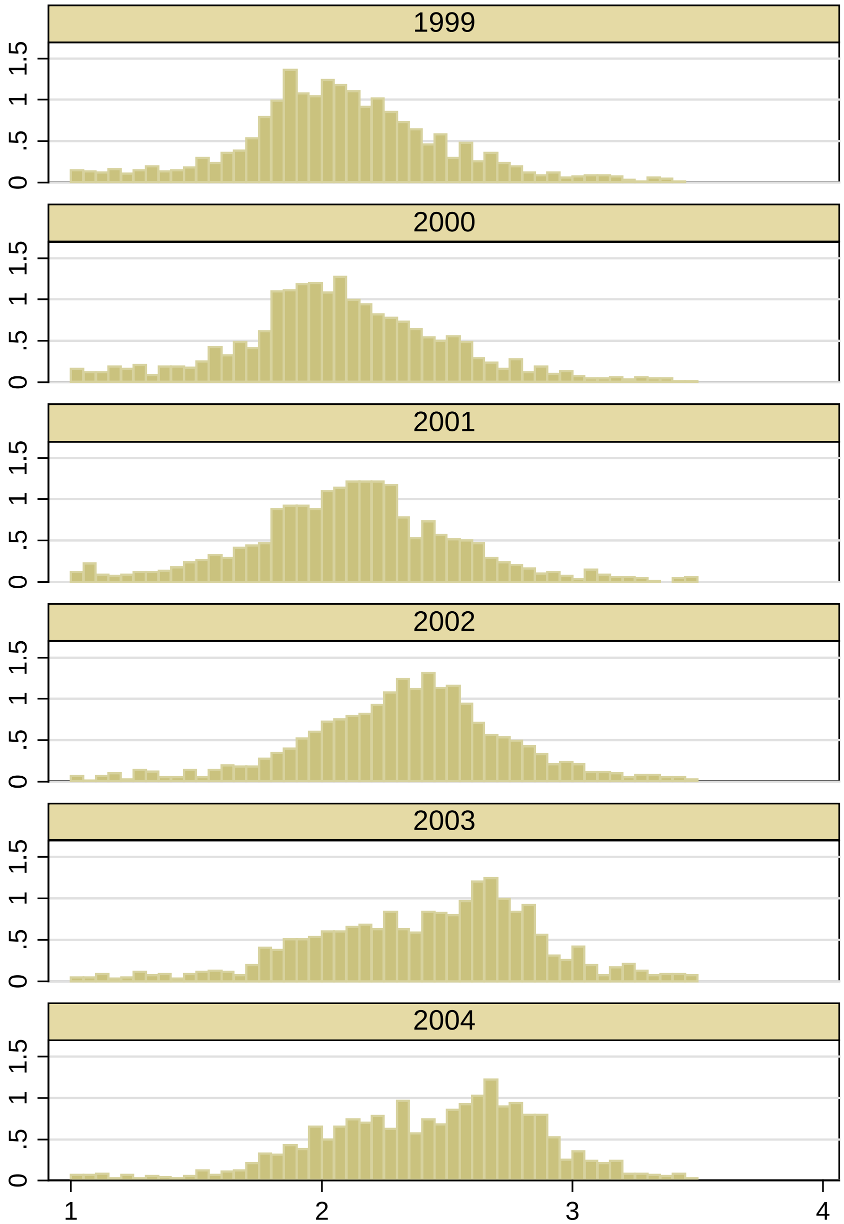
Nyman, John A., “Excess Demand, the Percentage of Medicaid Patients, and the Quality of Nursing Home Care,” *Journal of Human Resources*, 23 (1988), 76-92.

Scanlon, William J., “Nursing Workforce: Recruitment and Retention of Nurses and Nurse Aides is a Growing Concern,” U.S. General Accounting Office, GAO-01-750T (2001).

Weiss, Andrew, *Efficiency Wages: Models of Unemployment, Layoffs, and Wage Dispersions* (Princeton, Princeton University Press, 1990).

Wiener, J. M., M. P. Freiman, and D. Brown, “Nursing Home Care Quality: Twenty Years After The Omnibus Budget Reconciliation Act of 1987.” Commissioned by the Kaiser Family Foundation, (2007).

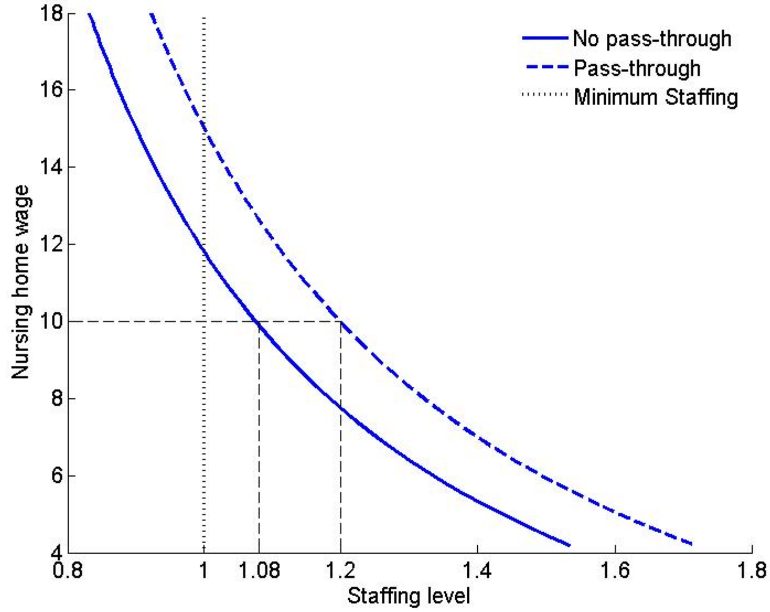
Figure 1. Distribution of Certified Nurse Aid (CNA) hours per patient day in Florida, New York, and Massachusetts Combined (1999-2004)



Notes: Florida, New York, and Massachusetts implemented the pass-through policy in 2002. Data is from the Online Survey Certification and Reporting (OSCAR) for years 1999 to 2004.

Figure 2. Labor Demand of Unconstrained and Constrained Nursing Homes

A. Unconstrained Nursing Home (l_{min} = 1)



l

B. Constrained Nursing Home (l_{min} = 1.2)

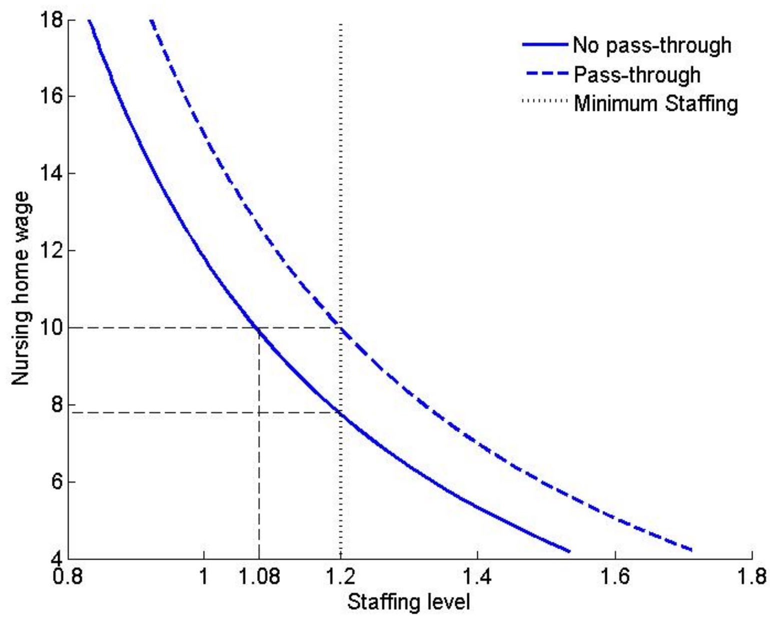
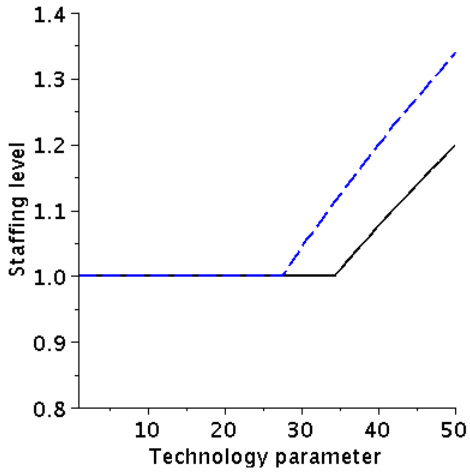
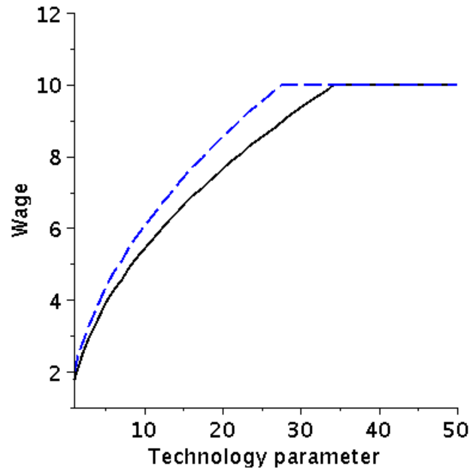


Figure 3. Impact of Pass-through by Technology Level

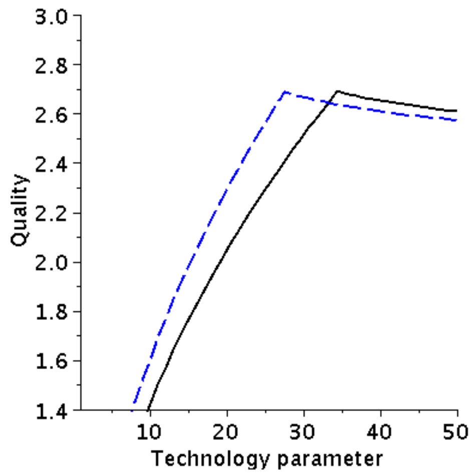
A. Staffing Level



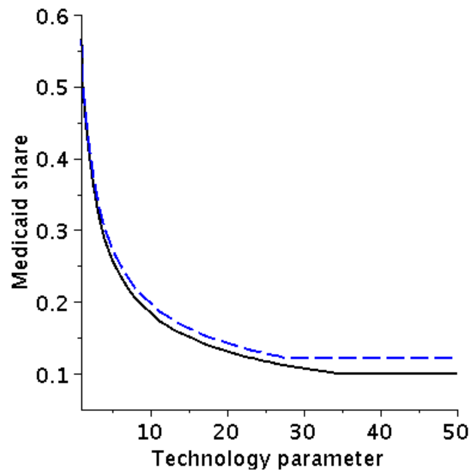
B. Wage



C. Quality of Care



D. Share of Medicaid Patients

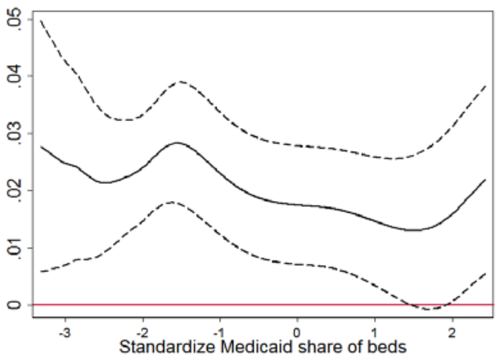


— No Pass-through - - Pass-through

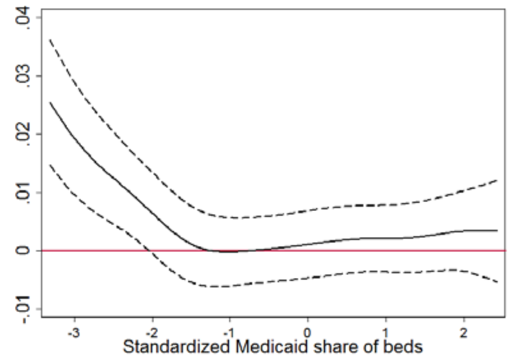
Figure 4. Non-parametric estimates of the impact of pass-through and Medicaid reimbursement rate over the Medicaid share of beds

I. Dependent variable : Log CNA hours

A. Impact of pass-through policy

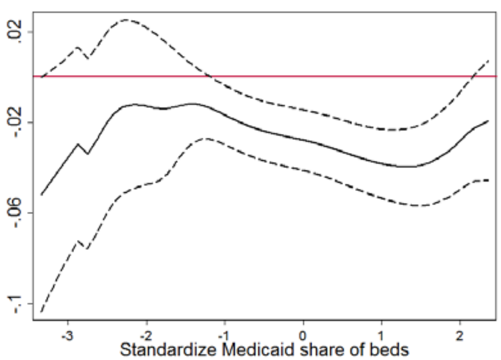


B. Impact of Medicaid reimbursement rate

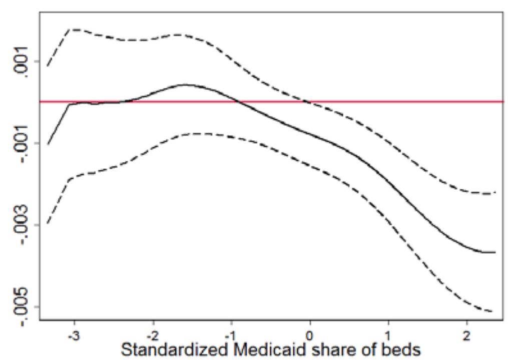


II. Dependent variable : Quality index

C. Impact of pass-through policy



D. Impact of Medicaid reimbursement rate



Notes: Dashed lines indicate standard error bands. Results are non-parametric estimates using a kernel density with 0.3 bandwidth.

Table 1. Efficiency Gains of the Two-Step FGLS procedure

Procedure		A: N=50, T=6, AR(1), ρ=0.8		B: N=50, T=12 AR(1), ρ=0.8		C: N=50,T=12,AR(2) ρ_1=0.6, ρ_2=0.3	
		Mean	SD	Mean	SD	Mean	SD
I. Facility fixed effect	β	1.090	0.761	0.964	0.532	0.960	0.508
w/ clustered SE	SE of β	0.772	0.147	0.533	0.0803	0.485	0.0770
II. Two-step DD	β	1.088	0.597	0.945	0.458	0.913	0.427
w/ clustered SE	SE of β	0.699	0.102	0.484	0.0577	0.431	0.0546
III. Two-step FGLS	β	1.055	0.423	0.994	0.306	0.925	0.336
w/ clustered SE	SE of β	0.473	0.0626	0.294	0.0328	0.341	0.0385
IV. Two-sted FGLS	β	1.055	0.423	0.993	0.306	0.925	0.336
w/ jackknife SE	SE of β	0.475	0.0408	0.205	0.0161	0.249	0.0222
Efficiency Gain		39%		45%		30%	

Notes: Simulations are for 50 states, with 0-100 facilities per state, over 6 or 12 time periods. We set the true β equal to one and carry out the exercise 100 times implementing (I) DD estimation with standard errors clustered at the state level, (II) Two-step DD with standard errors clustered by states, (III) Two-step FGLS with standard errors clustered by states, and (IV) Two-step FGLS with standard errors jack-knifed at the state level. The underlying AR process for the first two exercises are AR(1) and the last is AR(2). Efficiency gain calculates percentage reduction in the standard error of β from implementing procedure III relative to procedure I. SD and SE denote standard deviation and standard error, respectively.

Table 2. Summary Statistics

	Mean	SD	Min	Max	Obs
Dependent variable					
CNA hours per resident day	2.17	0.86	0	12	53135
State policies					
Wage pass-through:	0.25	0.44	0	1	53135
Medicaid payment rate (2004 $)	125.69	25.68	73.03	189.81	53135
Case-mix reimbursement	0.63	0.48	0	1	53135
Facility characteristics					
Acuity Index	11.08	1.57	3	22.75	53135
Nurse Practitioner/Physician Assistant	0.26	0.44	0	1	53135
Alzheimer's unit	0.19	0.39	0	1	53135
Percent private pay	24.83	20.33	0	100	53135
Percent Medicare	10.70	12.21	0	100	53135
Market (county) conditions					
Average # empty beds per nursing home	15.86	7.54	0	154	53135
RNs per hospital bed	1.46	0.42	0.03	6.58	52213
LPNs per hospital bed	0.18	0.10	0	1.33	52213
Nursing home beds per 1000 elders	51.03	17.70	2.20	364.67	53135
Managed care penetration rate (%)	16.82	14.46	0	55.32	53135
Area wage index	1.02	0.15	0.77	1.52	53135
Patient outcome variables					
Activities of Daily Living (ADL) decline	0.12	0.07	0	0.87	200411
Physical restraint	0.08	0.09	0	1.00	220690
Pressure ulcer worsening	0.06	0.04	0	0.45	213693
Persistent pain	0.07	0.06	0	0.79	211817
Quality index	0.0036	0.76	-1.644	7.193	200389

Notes: All variables, except for patient outcome variables are annual variables. Patient outcome variables are collected each quarter. Data comes from multiple sources including the Online Survey Certification and Reporting, Minimum Data Sets, Area Resource Files, and surveys of state Medicaid offices.

Table 3. Correlation of First-differenced Patient Outcome Variables

A_t	ADL	ADL	ADL	PR	PR	PU
B_t	PR	PU	PP	PU	PP	PP
Corr(ΔA,ΔB)	0.0091	0.1290	-0.0217	0.0031	0.0046	0.0008
Standard Error of Corr(ΔA,ΔB)	0.0023	0.0023	0.0023	0.0022	0.0022	0.0022

Notes: ADL=Activities of Daily Living decline rate, PR=Physical restraint rate, PU=Pressure ulcer worsening rate, PP=Persistent pain rate. $\Delta A = A_t - A_{t-1}$ and $\Delta B = B_t - B_{t-1}$.

Table 4. Factor Analysis on First-differenced Patient Outcome Variables

	A. One Factor Case	B. Two Factor Case	
	Loading on Factor1 (Eigenvalue=0.146)	Loading on Factor1 (Eigenvalue=0.148)	Loading on Factor2 (Eigenvalue=0.004)
Change in ADL decline rate	0.270	0.272	-0.003
Change in restraint rate		0.023	0.038
Change in pressure ulcer worsening rate	0.270	0.268	0.007
Change in persistent pain rate		-0.035	0.053

Table 5. Mean Values of the Quality Index by State

State	Mean quality	State	Mean quality	State	Mean quality
North Dakota	-0.3517	Illinois	-0.0835	Wisconsin	0.0685
Idaho	-0.3054	Washington	-0.0730	Louisiana	0.0785
Kansas	-0.2043	Alabama	-0.0611	West Virginia	0.0861
New York	-0.1966	Arizona	-0.0490	South Dakota	0.0968
Montana	-0.1939	New Mexico	-0.0474	Maryland	0.1171
Minnesota	-0.1858	Nebraska	-0.0426	Kentucky	0.1409
Utah	-0.1781	Virginia	-0.0384	Oklahoma	0.1710
Mississippi	-0.1504	Arkansas	-0.0308	Massachusetts	0.1754
New Hampshire	-0.1474	Texas	-0.0005	Connecticut	0.1795
South Carolina	-0.1463	Tennessee	0.0002	Delaware	0.1864
Iowa	-0.1419	Colorado	0.0005	Rhode Island	0.1883
Ohio	-0.1336	Florida	0.0157	Nevada	0.2003
Oregon	-0.1295	North Carolina	0.0218	New Jersey	0.2302
California	-0.1273	Geogia	0.0411	Pennsylvania	0.2421
Missouri	-0.1056	Vermont	0.0463	Wyoming	0.2904
Maine	-0.1029	Michigan	0.0503	Indiana	0.3482

Table 6. Duration Analysis on the Policy Variable

Dependent variable:	Pass-through policy			
	I	II	III	IV
log CNA hours (1999)	0.311	0.404		0.353
	(3.050)	(4.223)		(4.277)
Average quality (1999)	-0.271		-0.210	-0.198
	(1.277)		(1.577)	(1.605)
Medicaid payment rate (1999)	-0.0157	-0.0702	-0.0782	-0.0802
	(0.141)	(0.215)	(0.239)	(0.237)
Case-mix reimbursement (1999)	0.727	0.772	0.733	0.736
	(0.523)	(0.549)	(0.591)	(0.592)
Dummy: 2 years later	-0.555	-0.562	-0.551	-0.555
	(0.726)	(0.726)	(0.713)	(0.703)
Dummy: 3 years later	0.118	0.158	0.163	0.162
	(0.631)	(0.653)	(0.645)	(0.646)
Dummy: 4 years later	-0.711	-0.702	-0.700	-0.701
	(0.844)	(0.846)	(0.844)	(0.844)
Dummy: 5 years later	-1.367	-1.333	-1.330	-1.331
	(1.108)	(1.111)	(1.108)	(1.109)
Facility and market covariates	No	Yes	Yes	Yes
No. of observations	165	165	165	165

Notes: Initial policy, facility, and market covariates are 1999 averages. Excluded dummy is "one year later". Robust standard errors are in parentheses.

Table 7. DD and FGLS Estimation Results on the Full Sample

	Standard DID I	Two-step DID II	Two-step FGLS III	Two-step FGLS IV
Dependent variable: Log CNA hours				
Wage pass-through	0.0253	0.0192	0.0176*	0.0176*
	(0.0236)	(0.0131)	(0.00877)	(0.00903)
Medicaid payment rate	0.0149*	0.00390	0.00184	0.00184
	(0.00779)	(0.00776)	(0.00522)	(0.00575)
Case-mix reimbursement	-0.00900	0.000201	0.0146	0.0146
	(0.0173)	(0.0238)	(0.0181)	(0.0203)
AR(1) coefficient			0.79	0.79
No. of observations	52,213	288	288	288
Dependent variable: Quality Index				
Wage pass-through	0.0295	-0.0254	-0.0294**	-0.0294*
	(0.0290)	(0.0243)	(0.0143)	(0.0148)
Medicaid payment rate	-0.00256**	-0.00249**	-0.00115	-0.00115
	(0.00122)	(0.000965)	(0.000859)	(0.000947)
Case-mix reimbursement	0.0453	0.00826	0.0391	0.0391
	(0.0539)	(0.0529)	(0.0305)	(0.0331)
AR(1) coefficient			0.88	0.88
No. of observations	181,776	288	288	288
Standard Errors				
Clustered SE	O	O	O	-
Jack-knife SE	-	-	-	O

Notes: Staffing is annual data and quality is quarterly data. All specifications include facility, state, and year fixed effects. Facility and market variables are included in the standard DD procedure and the first-step of all two-step procedures. Standard error clustering and jack-knife are done at the state level. Medicaid repayment rate is in $10 increments. * and ** indicate significance at the 10% and 5% level respectively.

Table 8. DD and FGLS Estimation Results on Samples Stratified into Quartiles by Average Quality

	1st Quartile			2nd Quartile			3rd Quartile			4th Quartile		
	Standard DID	Two-step DID	Two-step FGLS	Standard DID	Two-step DID	Two-step FGLS	Standard DID	Two-step DID	Two-step FGLS	Standard DID	Two-step DID	Two-step FGLS
	I	II	III	I	II	III	I	II	III	I	II	III
Dependent variable: Log CNA hours												
Wage pass-through	0.0256	0.0318	0.0285	0.0234	0.0353**	0.0438**	0.0204	-0.0115	-0.0134	0.0318	0.0293	0.0213
	(0.0206)	(0.0213)	(0.0173)	(0.0242)	(0.0173)	(0.0189)	(0.0254)	(0.0220)	(0.0203)	(0.0259)	(0.0202)	(0.0177)
Medicaid repayment rate	0.0177	0.00299	0.00230	0.0140*	0.00737	0.00662*	0.0133	0.00228	0.00177	0.0160*	0.00772	0.00619
	(0.0114)	(0.0168)	(0.0108)	(0.00766)	(0.00449)	(0.00387)	(0.00827)	(0.00556)	(0.00502)	(0.00934)	(0.00771)	(0.00600)
Case-mix reimbursement	0.0144	0.0161	0.0339	-0.00560	0.00313	-0.00516	-0.0122	-0.00319	0.00134	-0.0271	-0.00280	0.0146
	(0.0335)	(0.0424)	(0.0255)	(0.0220)	(0.0257)	(0.0230)	(0.0170)	(0.0225)	(0.0204)	(0.0240)	(0.0345)	(0.0356)
AR(1) coefficient			0.86			0.23			0.15			0.4
No. of obs.	12,249	288	288	13,274	282	282	13,543	288	288	13,145	288	288
Dependent variable: Quality Index												
Wage pass-through	0.0132	-0.0267	-0.00859	0.0501	-0.0340	-0.0519	0.0245	-0.0211	-0.0274	0.0273	-0.0375	-0.0360
	(0.0273)	(0.0307)	(0.0248)	(0.0319)	(0.0459)	(0.0381)	(0.0343)	(0.0281)	(0.0169)	(0.0268)	(0.0353)	(0.0242)
Medicaid repayment rate	-0.00179	7.12e-05	0.000747	-0.00316**	-0.00194	-0.000783	-0.00255*	-0.00337**	-0.00282**	-0.00270**	-0.00333*	-0.00251
	(0.00143)	(0.00201)	(0.00168)	(0.00151)	(0.00134)	(0.00108)	(0.00135)	(0.00135)	(0.00130)	(0.00108)	(0.00170)	(0.00152)
Case-mix reimbursement	0.00779	0.0404	0.0670	0.0740	0.0328	0.0358	0.0445	-0.0731	-0.0119	0.0395	0.0291	0.0439
	(0.0494)	(0.0546)	(0.0512)	(0.0696)	(0.0665)	(0.0563)	(0.0706)	(0.0872)	(0.0551)	(0.0474)	(0.0773)	(0.0499)
AR(1) coefficient			0.6			0.46			0.6			0.55
No. of obs.	38,567	288	288	48,776	288	288	49,660	288	288	44,769	288	288

Notes: All specifications include facility, state, and year fixed effects. Facility and market variables are included in the standard DD procedure and the first-step of all two-step procedures. Standard errors are clustering at the state level for all procedures. Medicaid repayment rate is in $10 increments. * and ** indicate significance at the 10% and 5% level respectively.

Appendix. ***Efficiency Gains of FGLS relative to OLS***

Denote $E(U'U) = \Omega = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $A = E(X'X)^{-1}E(X'\Omega X)E(X'X)^{-1}$ and

$B = E(X'\Omega^{-1}X)^{-1}$ where A is variance of β under clustered OLS and B is the variance of β under FGLS. Let $Ex_1{}^2 = Ex_2{}^2 = \sigma_x{}^2$ where the subscripts on x indicate the time periods T=1, 2. Then $Ex_1x_2 = \kappa \cdot \sigma_x{}^2$. Substituting these into A and B we get $A = (1+\kappa\rho)(2\sigma_x{}^2)^{-1}$ and $B = (1-\rho^2)(2\sigma_x{}^2(1-\kappa\rho))^{-1}$. Calculating the percentage reduction by $(A-B)/A$ returns $1 - \dfrac{1-\rho^2}{1-\kappa^2\rho^2}$.