The Effect of Deregionalization on Health Outcomes: Evidence from Neonatal Intensive Care*

Seth M. Freedman†

RWJ Scholars in Health Policy Research Program,

University of Michigan

and

School of Public and Environmental Affars,

Indiana University

November 5, 2010

Abstract

The number of neonatal intensive care units (NICUs) in smaller, community hospitals increased greatly during the 1980s and 1990s, attracting deliveries away from hospitals with the most sophisticated NICUs. In this paper I estimate the causal effect on mortality of the level of care available at the hospital in which a very low birth weight infant is born. I exploit exogenous variation in distance from a mother's residence to hospitals offering each level. In contrast to previous estimates, these instrumental variables estimates suggest no differences in outcomes by level of neonatal care at the delivery hospital.

*Acknowledgments: I am particularly indebted to Judy Hellerstein, Ginger Jin, and Melissa Kearney for their advice and support throughout this project. I thank John Chao, John Cawley, Mark Duggan, Bill Evans, Craig Garthwaite, Darrell Gaskin, John Ham, Mara Lederman, Soohyung Lee, Tim Moore, Edward Norton, and seminar participants for helpful comments and suggestions. I am also grateful to Ciaran Phibbs for sharing level of care data, to OSHPD for assistance with the inpatient data, and to Bill Evans, Mark Duggan, Judy Hellerstein, and the University of Maryland Department of Economics for financial support in purchasing data. This work was partially supported by Agency for Healthcare Research and Quality Dissertation Fellowship Grant 1R36HS018266-01. The content of this work does not represent the views of the Agency for Healthcare Research and Quality or OSHPD. All errors are my own.

†Contact Information: 1415 Washington Heights, Ann Arbor, MI 48109. Email: sfreedm@umich.edu. Phone: (734) 936-1297. Fax: (734) 764-4338.

1 Introduction

Technological innovations over the past half century have greatly changed medical care for sick infants. Over this time Neonatal Intensive Care Units (NICU) have been developed to administer treatments such as mechanical ventilation, artificial surfactant, and extracorporeal membrane oxygenation (ECMO)[1] to sick, preterm, and underweight infants, and they have clearly lead to improved outcomes for these groups. For example, the 28-day mortality rate for infants weighing 1,000 to 1,499 grams (2.2 and 3.3 pounds) dropped from 52.2% to 6.7% between 1960 and 1990 (David M. Cutler & Ellen Meara 2000).[2]

Despite these long run gains, there is concern that NICUs have not diffused optimally as there has been a trend towards "deregionalization." The 1980s and 1990s saw a large increase in the number of NICUs in smaller, community hospitals that provide less sophisticated care compared to the original NICUs in large, regional hospitals (e.g. Marie C. McCormick & Douglas K. Richardson 1995, Rachel M. Schwartz 1996, Rachel M. Schwartz, Russell Kellogg & Janet H. Muri 2000). However, previous studies have found higher mortality rates for infants born in hospitals with these Community NICUs compared to those born in hospitals with Regional NICUs, conditional on observable demographic and health characteristics (e.g. Javier Cifuentes, Janet Bronstein, Ciaran S. Phibbs, Roderic H. Phibbs, Susan K. Schmitt & Waldemar A. Carlo 2002, Ciaran S. Phibbs, Laurence C. Baker, Aaron B. Caughey, Beate Danielsen, Susan K. Schmitt & Roderic H. Phibbs 2007, Ciaran S. Phibbs, Janet M. Bronstein, Eric Buxton & Roderic H. Phibbs 1996). Based on this evidence, organizations such as the March of Dimes and the American Academy of Pediatrics have advocated for a stronger regional system where high-risk mothers are referred to hospitals with Regional

[1]Mechanical ventilation assists infants whose lungs have not fully developed to breath. Artificial surfactant treats respiratory distress syndrome by helping the lungs to develop. ECMO machines pump blood out of the infant, oxygenate it, and pump it back into the infant if the infant's heart and lungs are too weak to oxygenate the blood on its own.

[2]I do not focus on costs in this paper, but nationwide it is estimated that medical care services for high-risk infants cost $16.9 billion in 2005 (`http://www.marchofdimes.com/peristats/slidesets/slideset_6_99.ppt`, last accessed on October 6, 2009). In the long run Cutler & Meara (2000) calculate a 510% rate of return to spending on infant health care between 1960 and 1990, accounting for the value of both lives saved and quality of life for surviving infants.

NICUs prior to delivery in order to minimize mortality.

This paper seeks to estimate the causal effect on mortality of the level of care available at the hospital in which a very low birth weight (VLBW) infant – under 1,500 grams or 3.3 pounds – is born. As an empirical matter, it is not clear that the worse outcomes experienced by infants born in hospitals with lower level NICUs are attributable to the hospital type per se. Even conditional on observable characteristics, infants born in different hospitals may have different underlying risk factors. Depending on the mechanisms behind any unobserved selection, conventional estimates of mortality differences by level of care could be biased in either direction. If infants born in hospitals with lower level NICUs have *lower* underlying mortality risk than those born in Regional NICUs, previous estimates will have *understated* the mortality penalty associated with being born in lower level hospitals. Alternatively, if infants born in hospitals with lower level NICUs have *higher* underlying risk factors, previous estimates will have *overstated* the mortality differences. Any bias implies the system of deregionalization might actually be more harmful or less harmful to mortality than currently believed. While deregionalization may affect many factors other than mortality, understanding the causal effect of level of care on mortality of high-risk infants is of first-order importance to making policy decisions about the organization of neonatal care.

I propose an instrumental variables strategy similar to that used by Mark McClellan, Barbara J. McNeil & Joseph P. Newhouse (1994) and David M. Cutler (2007) in their studies of heart attack treatment to overcome selection issues associated with a mother's choice of hospital. I exploit the distance a mother must travel to the nearest hospital of each level of care as a source of quasi-experimental variation in the type of hospital chosen. Distance is an important determinant of hospital choice for many medical treatments such as cardiac and cancer surgery (e.g. Cutler 2007, Daniel P. Kessler & Mark B. McClellan 2000, Mark McClellan & Joseph P. Newhouse 1997, Abigail Tay 2003) and for expectant mothers as well (Ciaran S. Phibbs, David H. Mark, Harold S. Luft, Deborah J. Peltzman-

Rennie, Deborah W. Garnick, Erik Lichtenberg & Stephen J. McPhee 1993). I also provide evidence that distance is likely to be exogenous to unobserved health outcomes in my data set, which is not surprising given evidence that NICU location is not correlated with the geographic variation in underlying infant health conditions (David C. Goodman, Elliott S. Fisher, George A. Little, Therese A. Stukel & Chiang hua Chang 2001). Using detailed data on all California VLBW births in urban and suburban areas between 1991 and 2001, I estimate how the birth hospital's level of care causally effects VLBW mortality.

My ordinary least squares (OLS) analysis yields estimates of 7.6%, 13.4%, and 31.8% higher risk-adjusted mortality rates for infants born at hospitals offering three lower levels of care relative to those born in hospitals offering the highest level of care. These results are consistent with the previous literature, but my instrumental variables estimates provide evidence that these OLS estimates are biased upward. The instrumental variables estimates are bounded well below the OLS estimates and are not statistically different from zero. My results are robust to including zip code of residence level controls, such as population density and racial characteristics, or zip code of residence fixed effects.

Comparing the OLS and the instrumental variable estimates reveals that infants born in hospitals with lower levels of care are negatively selected. This selection could occur if, for example, more uninformed mothers choose lower levels of care and have unobservably less healthy infants. This finding implies that relocating births to Regional NICU hospitals prior to delivery would not lead to lower mortality rates because the relocated infants would have higher unobserved mortality risk. In terms of mortality, deregionalization does not appear to have caused worse outcomes for high-risk infants.

While my results indicate that mortality does not differ by level of care at the hospital in which an infant is *born*, they do not imply that Regional NICUs are of no value. In fact, I show evidence that infants born in hospitals with the lowest levels of care are likely to be transferred to Regional NICU hospitals after birth, and the geographic distribution of hospitals does not impact the probability of transfer.[3] It is difficult to compare outcomes to the counterfactual world that experiences deregionalization but does not allow for post-birth transfer, but my findings suggest that mortality is not causally affected by the level of care at the birth hospital because high-risk infants eventually receive care in higher level hospitals if necessary.

It is also possible that the instrumental variable estimates represent a local average treatment effect. I find that my estimates are not heterogeneous across demographic subsamples, but there still may be heterogeneous effects along unobservable dimensions. If this is the case, instrumental variables would estimate the effect of level of care on mortality for an unobserved subgroup of infants whose mothers' choices of level of care are affected by the distance instruments. However, because variation in the instruments is directly related to deregionalization, any local effect is precisely the policy relevant effect. My estimates would imply that infants of mothers who choose to give birth in hospitals with lower level NICUs because these NICUs are available – the marginal group of infants whose delivery hospitals are impacted by deregionalization – do not experience higher mortality rates.

2 Background

2.1 Diffusion of Neonatal Intensive Care

As neonatal intensive care developed in the 1970s, few doctors and nurses were trained in neonatology. As a result, specialists were located in regional care centers, typically associated with large teaching hospitals. In 1976 a March of Dimes report recommended that hospitals offering delivery services be classified into three categories with the lowest providing no intensive care, and the highest providing the most complex care and acting as regional referral centers for high-risk mothers and infants (Committee on Perinatal Health 1976).

[3]My analysis sample does not include infants whose mothers reside in rural zip codes, so my results do not speak to issues regarding access to hospitals of varying levels of care or transfer between hospitals in rural areas.

Despite the March of Dimes' recommendations of a regionalized system, exactly the opposite began to occur over the 1980s and 1990s: there was a drastic increase in the number of NICUs, and many of the new entrants were smaller units in community hospitals (e.g. McCormick & Richardson 1995, Schwartz 1996, Schwartz, Kellogg & Muri 2000). While births increased by 17.6% between 1980 and 1995 in Metropolitan Statistical Areas (MSA), the number of hospitals with NICU beds doubled, the number of NICU beds more than doubled, and the number of neonatologists more than tripled (Embry M. Howell, Douglas Richardson, Paul Ginsburg & Barbara Foot 2002).[4] Additionally, American Hospital Association data reveal that 89% of the new NICUs that opened between 1980 and 1996 were lower level NICUs, as opposed to only 46% of the units established before 1980 (Laurence C. Baker & Ciaran S. Phibbs 2002).

Corinna A. Haberland, Ciaran S. Phibbs & Laurence C. Baker (2006) document that new lower level NICUs have in fact shifted deliveries of high-risk infants from Regional hospitals to the lower level hospitals in California. They find that becoming closer to a mid-level NICU, as a result of a new unit opening near a mother's residence, increases the probability that a very low birth weight infant is born in a hospital with a mid-level NICU by 17 percentage points and decreases the probability of being born in a hospital with a Regional NICU by 15 percentage points. Based on evidence that mortality rates are higher for infants born in hospitals with lower level NICUs, discussed in detail in below, and the course of deregionlization, the March of Dimes reaffirmed its recommendations in 1993 (Committee on Perinatal Health 1993). The American Academy of Pediatrics provided similar recommendations for more regionalized care in 2004 including recommendations for consistent definitions of care levels and the need for high-risk infants to be born in higher level facilities. (Committee on Fetus and Newborn 2004).

[4]Improving quality of care over time did lead to more infants surviving and spending longer periods of time in the NICU; however, Howell et al. (2002) calculate that by 1995 the number of available NICU bed-days exceeded medically necessary bed-days by a factor of 2.5.

2.2 Previous Estimates of Mortality Differences by Level of Care

Multiple authors have estimated how risk-adjusted mortality varies by level of neonatal intensive care at the hospital of birth, and many of these studies use the same California inpatient data set as this paper. The typical methodology includes a logistic regression of mortality on level of care indicators, controlling for demographic characteristics and health status. The specific results depend on the precise categorization of hospitals, but in general these studies find higher mortality as level of care decreases for groups of high-risk infants that NICUs are intended to care for. Phibbs et al. (1996) find that VLBW infants born in hospitals with the largest Regional NICUs have statistically lower mortality rates than the lower categories, but the lower categories, including hospitals with no NICU, do not differ from each other. Cifuentes et al. (2002) analyze a population of infants below 2,000 grams (4.4 pounds) and find that all levels except for the largest Community NICUs have higher mortality rates than Regional NICUs. As they restrict their sample to lower birth weight groups, the gradient becomes steeper. In the most recent study on the relationship between level of care and mortality, Phibbs et al. (2007) distinguish mortality rates by very narrow level and volume interactions. While not necessarily statistically significant within each level, they find decreasing mortality across levels and by volume within levels. Based on their estimates, they conclude that if 90% of VLBW deliveries in California urban areas had been relocated to hospitals with the largest Regional NICUs, 21% of VLBW deaths in 2000 could have been avoided.[5]

However, while high-quality hospital inpatient data sets allow the ability to control for many important covariates, mothers may select into different delivery hospitals based on characteristics not observed in the data. Such unobserved selection would lead to biased estimates of the mortality differences by level of care, and the direction of the bias would depend on the direction of the selection.

[5] They calculate this number only considering the sample of infants for whom they deem relocation geographically feasible and note that such relocation would require new large NICUs and the closure of some smaller NICUs.

One typical form of selection that biases estimates of the effect of health treatments on outcomes is selective referral. If mothers and physicians have additional information about the mother's health status, and higher risk mothers are referred to hospitals with Regional NICUs, mothers would be positively selected into lower levels of care. Therefore, the mortality differences relative to Regional NICUs would be underestimated. On the other hand, if mothers negatively select into lower levels of care over hospitals with Regional NICUs, the mortality differences would be overestimated. This case might arise if more uninformed mothers are more likely to choose hospitals with lower levels of care over hospitals with Regional NICUs, and if infants of these uninformed mothers have higher unobserved mortality risk.

2.3 Marginal Effects of Medical Care

This paper is also related to the health economics literature that attempts to determine the marginal effects of medical treatments and technology. As with neonatal care, time series evidence suggests that most new technologies have led to vast improvements in health outcomes over time and the monetized benefits have outweighed the costs (David M. Cutler & Mark McClellan 2001, Robert E. Hall & Charles I. Jones 2007, Bryan R. Luce, Josephine Mauskopf, Frank A. Sloan, Jan Ostermann & L. Clark Paramore 2006, Kevin M. Murphy & Robert H. Topel 2003). However, comparisons of health care expenditures and outcomes across geographic regions have found that higher spending areas do not achieve better outcomes (Katherine Baicker, Kasey S. Buckles & Amitabh Chandra 2006, Katherine Baicker & Amitabh Chandra 2004, Elliott S. Fisher, David E. Wennberg, Therese A. Stukel, Daniel J. Gottlieb, F. L. Lucas & Etoile L. Pinder 2003*a*, Elliott S. Fisher, David E. Wennberg, Therese A. Stukel, Daniel J. Gottlieb, F. L. Lucas & Etoile L. Pinder 2003*b*, Victor R. Fuchs 2004). Given this contradiction, researchers have taken advantage of quasi-experimental variation to better compare individuals who differ only in their treatment and not in other unobserved dimensions to estimate causal effects of treatment. Here I highlight two portions of this literature that are most related to this paper: research on the effects of infant health care and research using a similar identification strategy to that used in this paper.

Studies of the returns to incremental units of infant health care find mixed results. Douglas Almond & Joseph J. Doyle (2008) exploit a California policy extending minimum length of hospital stays following delivery and the discontinuity in stay length for infants born just before and just after midnight. They find no effect of increased stay length on health outcomes for uncomplicated infants. William N. Evans, Craig Garthwaite & Heng Wei (2008) and William N. Evans & Craig L. Garthwaite (2009) exploit the same policy and find no effect of length of stay on readmission rates on average, but that longer length of stay does lead to reduced hospital readmission rates for infants with complications, pre-existing conditions, or higher predicted medical need. Using a regression discontinuity design, Douglas Almond, Joseph J. Doyle, Amanda E. Kowalski & Heidi Williams (2008) find that infants just below the VLBW cutoff receive more treatment and experience lower mortality rates than those just above the VLBW cutoff. Taken together, these studies imply that, at least for high-risk infants, increased treatment can be beneficial.

McClellan, McNeil & Newhouse (1994) and Cutler (2007) use a similar identification strategy to this paper's strategy in order to estimate the effect of catheterization and revascularization, respectively, following a heart attack on mortality. As with infant care, there are two selection concerns in this context, although the mechanisms are slightly different. First, the healthiest patients may have less need for these intensive surgeries. Second, the sickest patients may forego surgery due to a higher risk of dying during the procedure. To account for selection, both papers use differential distance to the nearest hospital providing surgery relative to the nearest non-surgery hospital as an instrument for whether a patient receives surgery.[6] Both studies find that instrumental variable estimates of the benefit of

[6]Other authors have also used distance as a source of exogenous variation to predict patient flows in order to estimate the effect of volume (Gautam Gowrisankaran, Vivian Ho & Robert J. Town 2006) and competition (Gautam Gowrisankaran & Robert J. Town 2003, Kessler & McClellan 2000, Tay 2003) on health outcomes.

intensive surgery are substantially lower than the ordinary least squares estimates, although Cutler (2007) finds that the monetized benefits still outweigh the costs. Amitabh Chandra & Douglas O. Staiger (2007) use the same identification strategy to examine how the returns to intensive heart attack treatment vary between more and less intensive geographic areas.

3 Data

3.1 Linked Birth Data

My empirical analysis requires detailed data describing infants' hospitalizations and outcomes. The primary data set I utilize is the Linked Patient Discharge Data/Birth Cohort File (LPDD/BCF) created by the California Office of Statewide Health Planning and Development (OSHPD). This data set includes records of all births in non-Federal hospitals in the state of California. I have obtained data files for the years 1991 to 2001, comprising approximately six million births. In addition to including observations of all births from a large state, the main advantage of this data set is that it links additional data to an infant's hospital discharge record. First, it links an infant's delivery hospital discharge record to the mother's discharge record and all subsequent records resulting from transfers or readmissions to California hospitals within the first year of life. For each hospitalization, the data set includes detailed diagnosis and treatment variables, summary variables such as length of stay and hospital charges, and patient information including zip code of residence. Second, the hospital discharge data are linked to vital statistics data on births and infant deaths within the first year of life, which include gestation, birth weight, number of prenatal care visits, month prenatal care began, and demographics, such as the mother and father's race and education. Additionally, these records provide information on infant mortality within the first year of life, even if death occurred outside of the hospital.

The VLBW analysis sample that I consider includes infants weighing between 500 and 1,500 grams (1.1 and 3.3 pounds) at birth. Of the initial 6.1 million birth observations with non-missing birth weight, 72,275 fall in this birth weight range.[7] To obtain my analysis sample, I first exclude observations with a missing zip code of residence, a zip code of residence outside the state of California, a missing hospital identification number, or that are delivered in a hospital without a delivery unit. The remaining sample contains 65,567 birth observations.

I then make three restrictions to maintain a sample that is as broad as possible but that excludes observations with an unusual hospital choice set. I first drop 2,704 observations where the mother's county of residence is "non-metro" according to the Office of Management and Budget.[8] This restriction excludes a small group of infants from the most rural areas for whom access to neonatal care is quite different from other residents of the state. Additionally, the previous literature has focused on deregionalization and the effect of level of care on outcomes in metropolitan areas (Howell et al. 2002, Phibbs et al. 2007) where policy recommendations about delivery relocation would be most feasible. Second, I drop 7,627 infants delivered in Kaiser owned hospitals. Mothers who choose a Kaiser hospital for delivery must be covered by Kaiser insurance, and mothers covered by Kaiser insurance must deliver in a Kaiser owned hospital; therefore, choice of hospital is restricted for this group.[9] Third, I exclude 4,113 observations diagnosed with a congenital anomaly. This restriction is consistent with the previous literature (Phibbs et al. 2007), and it also excludes observations most likely to be selectively referred to higher levels of care due to a diagnosis during the prenatal period. I also exclude 8,115 observations of fetal deaths, which are infants who die prior to delivery and, therefore, never receive neonatal care (Phibbs et al. 2007). Finally, because I cluster standard errors at the zip code level and estimate models with zip code fixed effects, I exclude 96 observations for which the mother's zip code of residence has no

[7] The full data set includes 6,221,001 births of which 1.54% of the observations have a missing birth weight.

[8] Based on 1993 USDA Rural-Urban Continuum Codes that are calculated from the 1990 Census. Source: `http://www.ers.usda.gov/briefing/rurality/ruralurbcon/priordescription.htm`.

[9] In my analysis sample, 88% of mothers with Kaiser coverage deliver in a Kaiser hospital, and 97% of mothers who deliver in a Kaiser hospital have Kaiser coverage. In results not shown here, regressions similar to the first stage regressions discussed below for the sample of Kaiser insured mothers show that distance has very little power in predicting the level of care chosen for delivery. This is in contrast to the strong predictive power of distance for the analysis sample discussed in Section 5.

other observations remaining in the data. In the online appendix, I show that my results are robust to each of these sample restrictions.

I choose my sample of high-risk infants using birth weight as the health proxy because it is the best measure of an infant's health stock at birth (Douglas Almond, Kenneth Y. Chay & David S. Lee 2005, Cutler & Meara 2000).[10] VLBW infants are the population most of interest because they contribute disproportionately to costs and mortality. Mean charges for my VLBW sample are $209,000, compared to $21,000 for low birth weight infants (1500 to 2500 grams or 3.3 to 5.5 pounds) and $2,630 for normal birth weight infants (above 2500 grams or 5.5 pounds). Likewise, length of stay after birth averages 50.6 days for VLBW infants, 9.2 days for low birth weight infants, and 3.0 days for normal birth weight infants. Additionally, VLBW infants make up the vast majority of infant mortality. The main outcome I focus on in this paper is neonatal mortality, defined as mortality within twenty-eight days of birth or within one year if an infant is continuously hospitalized since birth. VLBW infants have a neonatal mortality rate of 15.7%, compared to 0.7% for low birth weight infants and 0.1% for normal birth weight infants.

3.2 Hospital Data

My empirical analysis also requires data describing the level of neonatal care offered by each hospital that delivers infants. I obtain data from the authors of Phibbs et al. (2007) that differentiate hospitals into six levels of neonatal care based on the treatments each hospital provides in a given year. First, they use OSHPD hospital financial data to determine which hospitals have neonatal intensive care beds. Second, they use ICD-9 CM procedure codes in the hospital inpatient data to identify which hospitals perform particular procedures. As a guide, they define levels of care consistent with the six levels outlined in the American

[10] Relative to gestation, another summary of health status at birth, Almond et al. (2008) note that birth weight is more accurately recorded, less likely to be missing in the data, and less likely to be manipulated by delaying birth because it is not possible to know birth weight *ex ante*. Additionally, they find empirical evidence that the recording of birth weight is not manipulated by physicians.

Academy of Pediatrics 2004 report.[11] Table 1 lists the six levels and their corresponding procedures. Third, the authors confirmed level of care designations through conversations with hospital personnel.

I collapse these detailed categories into four levels of care, which I refer to as No NICU, Intermediate NICU, Community NICU, and Regional NICU hospitals. No NICU hospitals provide birthing services and well-baby care, but no neonatal intensive care (Level I in Table 1). Intermediate NICUs care for mildly ill infants but do not provide mechanical ventilation (Level II). Community NICUs include any unit that provides mechanical ventilation and either does not provide major surgery or provides surgery but treated less than 50 VLBW infants in 1991 (IIIA, small IIIB, and small IIIC).[12] Finally, Regional NICUs include those that provide major surgeries and treated greater than 50 VLBW infants in 1991, or any unit that provides cardiac bypass and/or ECMO, the two most specialized surgical procedures, regardless of size (large IIIB, large IIIC, and all IIID).

Table 2 shows the number of hospitals by level and year between 1991 and 2001. There are 161 No NICU, 58 Intermediate, 41 Community, and 36 Regional NICU hospitals at the beginning of my sample in 1991. The total number of hospitals providing any birthing services falls from 296 in 1991 to 269 in 2001. But, consistent with deregionalization, the number of Community NICUs increases from 35 to a peak of 57 in 1999. 10 hospitals open new NICUs at the Community level and 21 hospitals upgrade an Intermediate NICU to the Community level. As a result of these upgrades, the aggregate number of Intermediate NICUs actually decreases from 58 to 45 over the sample period; however, there are also 15 hospitals that open new NICUs at the Intermediate level. Not surprisingly, the number of Regional NICUs, the largest, most well established, and most expensive units, remains relatively constant over the sample period.

[11] The authors utilize the draft version of the American Academy of Pediatrics report because the final version does not include a category that provides unrestricted ventilation but no surgery, a level of service many CA units provide.

[12] I use the number of VLBW infants treated in 1991 to identify this classification to prevent hospitals from changing levels due to changes in demand during my sample period.

4 Empirical Framework

4.1 Baseline Model

I begin by estimating the average difference in mortality rates by level of care at the delivery hospital, controlling for observable characteristics of the mother and infant. This estimation strategy is comparable to the methodology of the previous literature and provides "risk-adjusted" mortality differences. The regression equation is as follows:

$$y_{izt} = \alpha + N_{izt}\beta^N + I_{izt}\beta^I + C_{izt}\beta^C + \mathbf{X}_{izt}\mathbf{\Gamma} + \varepsilon_{izt} \tag{1}$$

The unit of observation is infant i, whose mother resides in zip code z, born in year t. The dependent variable, y_{izt}, is a neonatal mortality indicator that is equal to one if an infant dies within 28 days of birth or within one year if continually hospitalized since birth, and zero otherwise.[13] $\mathbf{X}_{izt}$ is a vector of observable determinants of infant izt's health. These controls include time (year, month, and day of week indicators); mother's demographics such as age, race, ethnicity, and insurance coverage;[14] and health related controls such as the infant's sex, birth weight, and diagnoses.[15]

The three explanatory variables of interest, N_{izt}, I_{izt}, C_{izt}, are indicators equal to one if infant izt is born in a hospital with No NICU, an Intermediate NICU, or a Community NICU, respectively. Being born in a hospital with a Regional NICU is the excluded group, so the β^j coefficients are interpreted as the difference in mortality when born in a hospital with

[13]The online appendix shows that results are robust to measuring mortality outcomes over other time windows.

[14]Specific demographic controls are age, age squared, and indicators for black, other race, Hispanic, Medicaid, HMO, and self-pay.

[15]Specific health controls are parity, sex, multiple birth status, an indicator for having a clinical condition, indicators for small and large for gestational age, birth weight dummies at 100-gram increments, the number of prenatal care visits, and the month in which prenatal care began. The clinical condition indicator is equal to one for infants having at least one of the following conditions identified in Phibbs et al. (2007): hydrops due to isoimmunization, hemolytic disorders, fetal distress, fetus affected by maternal condition, oligohydramnios, other high-risk maternal conditions, placenta hemorrhage, premature rupture of membrane, and prolapsed cord. The online appendix shows that results are robust to specifiying birth weight in 50-gram increments, interacting birth weight dummies with sex, controlling for cesarean delivery, and including separate indicators for each clinical condition.

level of care j relative to being born in a Regional NICU hospital. For this specification to estimate the causal effect of level of care on mortality, hospital choice must be uncorrelated with unobserved determinants of mortality captured by the error term, ε_{izt}, conditional on the observable characteristics, $\mathbf{X}_{izt}$ ($E\left[\mathbf{H}'_{izt}\varepsilon_{izt}|\mathbf{X}_{izt}\right] = \mathbf{0}$, where $H_{izt} = [N_{izt}, I_{izt}, C_{izt}]$). Depending on the direction of any correlation between level of care and unobserved mortality, OLS estimates could overstate or understate the mortality differences.

Sample means by level of care in Table 3 show that there are clear unconditional differences in mortality rates by level of care at the hospital in which an infant is born. Neonatal mortality rates fall from 21.9% for VLBW infants born in No NICU hospitals, to 16.9% in Intermediate NICU hospitals, 15.5% in Community NICU hospitals, and 14.7% in Regional NICU hospitals. However, there are also differences in important observable characteristics. OLS regressions control for these observable characteristics, but these differences motivate the concern that there may be differences in unobservable dimensions as well (Joseph G. Altonji, Todd E. Elder & Christopher R. Taber 2005). Mothers' demographic characteristics differ by level of care, but not monotonically. For example, 9.8% of mothers giving birth in No NICU hospitals, 20.5% in Intermediate NICUs, 12.8% in Community NICUs, and 18.6% in Regional NICUs are black. The percentage of mothers covered by Medicaid and the percentage without any college education decreases substantially from No NICU, to Intermediate NICU, and to Community NICU hospitals, but the percentage in Regional NICU hospitals is higher than the percentage in Community NICU hospitals. These large differences indicate selection into level of care by mothers' demographics, but the direction of the selection is ambiguous.

There are also clear patterns of selection on infant health characteristics. Consistent with selection of healthier infants into lower levels of care, infants born at lower levels are less likely to be multiple births, have slightly higher birth weight and longer gestation, are less likely to have a clinical diagnosis, are less likely to be small or large for their gestational age, and experience lower hospital charges and shorter lengths of stay. Given the differences in observed characteristics by level of care, there are likely differences in unobserved characteristics as well. Therefore, accounting for non-random selection is important, though the direction of the bias is again unclear *ex ante*.

4.2 Estimating Causal Effects

Because OLS estimates may not be able to control for all determinants of mortality, I utilize instrumental variables to overcome unobserved selection. With three endogenous explanatory variables, at least three instruments are necessary to identify the empirical model. I construct three instruments based on the distance from a mother's residence to each level of care, which I define in more detail below. The first stage regressions are as follows:

$$\begin{aligned} N_{izt} &= \delta^N + \mathbf{D}_{zt}\mathbf{\Pi}^N + \mathbf{X}_{izt}\mathbf{\Gamma}^N + \mu_{izt}^N \\ I_{izt} &= \delta^I + \mathbf{D}_{zt}\mathbf{\Pi}^I + \mathbf{X}_{izt}\mathbf{\Gamma}^I + \mu_{izt}^I \\ C_{izt} &= \delta^C + \mathbf{D}_{zt}\mathbf{\Pi}^C + \mathbf{X}_{izt}\mathbf{\Gamma}^C + \mu_{izt}^C \end{aligned} \tag{2}$$

Notation is as above with $\mathbf{D}_{zt}$ representing a 3×1 vector of instrumentas, each $\mathbf{\Pi}^j$ representing a 1×3 vector of three first stage coefficients, and each μ_{izt}^j representing a first stage error term.

After estimating these first stage regression equations, Equation (1) is estimated via two stage least squares (2SLS).[16] These estimates of β^N, β^I, and β^C will be consistent if the instruments are uncorrelated with the error term in Equation (1) ($E[\mathbf{D}'_{zt}\varepsilon_{izt}|\mathbf{X}_{izt}] = \mathbf{0}$)

[16] Both the dependent variable and the endogenous explanatory variables in this model are binary. Jay Bhattacharya, Dana Goldman & Daniel McCaffrey (2006) point out that two stage least squares can lead to inconsistent estimates when the mean probability of the binary dependent variable is close to zero or one, or when there is more than one endogenous binary treatment variable. On the other hand, Joshua D. Angrist (2001) argues that linear models still provide good approximations of average causal effects, parameter estimates directly correspond to the relevant average treatment effects, and nonlinear models depend on the distributional assumptions and are inconsistent if these assumptions are incorrect. Jeffrey M. Wooldridge (2001) points out that some of the assumptions behind average treatment effects are not precisely true with binary outcomes, but linear methods may still produce reasonable average treatment effect estimates. I have estimated my OLS specifications with both probit and logit models and find marginal effects that are almost identical to the OLS coefficient estimates presented in Section 5. In Section 6 I discuss a robustness check to address concerns about 2SLS when some of the treatment variables have means near zero.

and are strong determinants of the type of hospital a mother chooses,[17] conditional on the other observable characteristics. Intuitively, identification of β^N, β^I, and β^C comes from comparing mortality for otherwise identical infants who are born at different levels of care because they live at different distances from each level of care. For example, β^C is identified from differences in mortality outcomes between infants who are and are not born in hospitals with Community NICUs because their mothers live within close or far proximity to a hospital offering a Community NICU. This comparison emphasizes the importance of the assumption that mothers living at different distances from each level of care not have infants that differ in unobserved determinants of mortality.

Since this strategy requires the location of NICUs to be exogenous to VLBW infant health, it is worth briefly explaining the process by which NICUs have diffused and mothers choose hospitals. Most importantly, diffusion has been driven by many factors unrelated to the health of VLBW infants. Over time the technologies and trained specialists necessary to operate NICUs became more prevalent, and therefore, NICU adoption became feasible for community hospitals. It has been hypothesized that so many hospitals adopted lower level NICUs in order to compete for profitable obstetric patients (McCormick & Richardson 1995). Ninety-seven percent of births are covered by private or public insurance (Rebecca B. Russell, Nancy S. Green, Claudia A. Steiner, Susan Meikle, Jennifer L. Howse, Karalee Poschman, Todd Dias, Lisa Potetz, Michael J. Davidoff, Karla Damus & Joann R. Petrini 2007), so most families are shielded from the full cost of infant care. One way for hospitals to compete for these patients is to invest in signals of quality, which might attract risk-averse mothers.[18] Hospitals are particularly motivated to attract obstetric paitents since mothers are typically

[17]More formally, it must be the case that the instruments are sufficiently linearly related to $\mathbf{H}_{izt}$ that $E[\mathbf{Z}'_{izt}\mathbf{H}_{izt}]$ is of full column rank, where $\mathbf{Z}_{zt} = [\mathbf{D}_{zt}, \mathbf{X}_{izt}]$. It is also necessary for the instruments to be sufficiently linearly independent so that $E[\mathbf{Z}'_{izt}\mathbf{Z}_{zt}]$ has full column rank (Wooldridge 2001).

[18]This type of competition is not unique to neonatal intensive care. Theoretically, non-price competition can potentially lead to over-provision of services known as a "medical arms race" (Martin Gaynor 2006). David Dranove, Mark Shanley & Carol Simon (1992) find that decreases in market concentration lead to increases in the number of hospitals offering various high tech services in that market. Others have shown that hospitals expand their capacity to perform certain procedures in order to deter other hospitals from adopting that procedure (Leemore S. Dafny 2005), and hospitals adopt particular technologies in order to steal business from their competitors (Philipp Schmidt-Dengler 2006).

young, healthy, and likely to return to the hospital for the later care of their families if they have a positive birth experience (Bernard Friedman, Kelly J. Devers, Claudia A. Steiner & Steven H. Fox 2002). Furthermore, NICUs themselves are profitable, because reimbursement rates are relatively high, and managed care has been hesitent to limit reimbursement of infant care (Jill R. Horwitz 2005, see online appendix).

A NICU is likely an effective tool for attracting patients of all risk levels. Most preterm labor is spontaneous.[19] Expectant mothers usually deliver in the hospital where their obstetrician has delivery privileges, so they in effect choose their delivery hospital when they choose their obstetrician early in their pregnancy. If risk-averse mothers plan ahead when choosing their obstetrician and delivery hospital, the presence of a NICU is likely to factor into their decision. A mother likely considers travel time, convenience for family members, perceived quality of care, and the possibility of transfer if higher quality care is needed. If utility is increasing in perceived quality of care and decreasing in travel time, a community hospital with a NICU can attract nearby mothers willing to trade additional perceived quality at a further Regional NICU in favor of the increased convenience of choosing the nearby hospital. Furthermore, if mothers tend to choose local obstetricians who are likely to have priveldges in local hospitals, mothers will be more likely to choose nearby hospitals. The motives to compete for deliveries of all risk levels, evidence that growth of neonatal resources has outpaced medical need, and findings that the location of neonatal intensive care resources are uncorrelated with markers of need such as occurrences of VLBW or preterm births (Goodman et al. 2001), support the exogeneity of NICU location to VLBW infant health. Below I provide further evidence from my data supporting this claim.

[19] In 50% of cases, doctors are not even able to determine the cause *ex post*. Forty to fifty percent of cases with an identified cause are traced to an infection, but often mothers show no signs of these infections prior to labor (`www.marchofdimes.com/peristats`, last accessed on September 29, 2009). There are a variety of documented correlates of preterm delivery such as tobacco use, nutrition, stress, and demographics but there is in fact little understanding of what conditions and events can be used to predict and diagnose preterm labor before it occurs (Richard E. Behrman & Adrienne Stith Butler 2007).

4.3 The Instruments

In this section, I describe how I calculate the three distance instruments and provide empirical evidence that they are exogenous to unobserved VLBW mortality. I first calculate the straight line distance from the center of each patient's zip code of residence to each hospital using GIS software. Hospital location is obtained from OSHPD's publicly available geocoded data of hospital latitude and longitude.[20] I then construct three instruments that represent the differential distance between the nearest hospital of a given level of care or higher and the nearest hospital with a Regional NICU, as follows:

$$NoDist_{zt} = D(Reg_{zt}) - min[D(No_{zt}), D(Int_{zt}), D(Com_{zt}), D(Reg_{zt})] \tag{3a}$$

$$IntDist_{zt} = D(Reg_{zt}) - min[D(Int_{zt}), D(Com_{zt}), D(Reg_{zt})] \tag{3b}$$

$$ComDist_{zt} = D(Reg_{zt}) - min[D(Com_{zt}), D(Reg_{zt})] \tag{3c}$$

The $D(\cdot)$ operator indicates the distance from zip code z at time t to the nearest hospital offering a particular level of care. These measures can be thought of as the number of miles *saved* by choosing the nearest hospital with at least a particular level of care over the nearest hospital with the highest level of care, and therefore get larger as an individual lives closer to a hospital offering the particular level of care.

When using differential distance, the hospital choice decision is modeled as a function of distance to each lower level of care relative to distance to Regional NICU hospitals.[21] It emphasizes the fact that mothers make a trade off when choosing a lower level of care at a closer hospital – they forego potentially higher quality care in exchange for a shorter travel time.[22] Also, these three measures will always take on values greater than or equal to

[20]OSHPD only provides this data for currently existing facilities. For those facilities for which I do not have exact location, I use the center of the hospital's zip code obtained in the OSHPD State Utilization File of Hospitals.

[21]Cutler (2007) and McClellan, McNeil & Newhouse (1994) also use differential distance as their instruments by subtracting distance to the nearest hospital from distance to the nearest hospital offering heart surgery.

[22]A model with four instruments based on distance to each of the four levels of care would achieve the

zero due to the $min[\cdot]$ operator in (3), and they equal zero if an individual lives closer to a Regional NICU than one of the lower levels of care. This specification captures the fact that if a hospital nearby offers a particular level of care, a mother can also receive lower level care by traveling to the same hospital.

These distances are not the only way one could specify exposure to NICUs. I utilize this method to best proxy for the cost of obtaining each level of care; although, one could also specify distance based on the distance to the nearest hospital of a specific level of care (instead of the nearest hospital with a particular level or higher). Other potential measures of exposure include hospital market shares or the number of hospitals of each level within a given radius. I choose distance so as not to impose potentially endogenous market definitions. It is also important to point out that under the assumptions of the first stage relationships in Equation (2) are reduced form equations and do not necessarily provide structural parameters of the neonatal intensive care level demand function. The goal is not to estimate structural parameters of hospital choice, but instead to exploit the exogenous variation in distance that directs patients to different levels of care.[23]

Table 4 provides summary statistics of the four distance measures used to construct the instruments and of the three instruments themselves. On average, mothers of VLBW infants in my sample live 3.7, 5.7, 8.1, and 14.8 miles from the nearest hospital offering any birthing services, at least Intermediate care, at least Community care, and Regional care, respectively. The average number of miles saved by traveling to the nearest hospital with no NICU or higher relative to the nearest Regional NICU is 11.2. The average number of

same goal, as it would condition on distance to the nearest Regional NICU in each first stage regression. Using differential distance is equivalent to including all four distance measures separately, but restricting the coefficient on the Regional distance variable. 2SLS results, not shown here, without this functional form assumption are almost identical to those presented in Section 5.

[23]Phibbs et al. (1993) estimates McFadden conditional logit models of hospital choice, and their model includes features such as distance from a mother's residence and presence of a neonatal intensive care unit. Additional work in this area is left to future research, as estimating such demand functions is important for understanding how mothers choose hospitals and why hospitals choose to provide various levels of care. Dana Goldman & John A. Romley (2008) find that Medicare pneumonia patients in Los Angeles place a high value on non-medical amenities when choosing a hospital for treatment. Amenities, such as private roos, jacuzzis, etc., may also be an important tool for hospitals competing for maternity patients.

miles saved traveling to the nearest hospital with at least an Intermediate NICU or at least a Community NICU is 9.1 and 6.8 miles, respectively. These measures have wide variation, each with standard deviations around 20 miles, or two to three times their means.

I now provide a set of summary statistics supporting the assumption that differential distance is uncorrelated with the error term in Equation (1). Table 5 lists sample means of observable characteristics by the three instruments. While there is no formal test of this assumptoin, if a detailed list of observable characteristics are independent of differential distance, it is likely to be the case that unobservable characteristics are as well (Altonji, Elder & Taber 2005). For each instrument the table shows sample means for three groups: those observations with zero differential distance and those with differential distance below and above the median, conditional on non-zero differential distance.

Other than the proportion of mothers who are black, which is about twice as large for observations at zero differential distance compared to individuals above the median for all three instruments, mothers' demographics show little variation by distance. For example, the percent of mothers covered by Medicaid ranges from 48.3% to 52.1% for the Community distance groups. In contrast, this figure had a gap of 13.6 percentage points between infants born in No NICU and Community NICU hospitals in Table 3. Most importantly, infant health characteristics do not differ much across distance groups. While the number of prenatal visits is slightly lower for those with zero miles saved, the month prenatal care began is similar across groups and there are no large differences in parity, multiple births, birth weight, or gestation.

To understand the result that the proportion of black mothers varies with differential distance, the bottom portion of the table presents means of zip code level characteristics. These variables are collected from the 2000 census and merged to the mother's zip code of residence, and I present means treating each birth as an observation. Here, there are some potentially important differences by differential distance as median household income increases, percent urban decreases, and population density decreases across columns for each distance variable. Consistent with the result for black mothers, this suggests that mothers living closer to lower level NICUs relative to Regional NICUs live in less densely populated and less urban areas.[24]

Despite these differences, the variation in differential distance is not driven by population density alone. Figure A.1 in the online appendix displays the geographic distribution of the community distance and the intermediate distance variables at their 1991 baseline, both for the whole state and zooming in on the five counties comprising the Los Angeles Metro area. In these maps, the lightest colored zip codes have no births in the VLBW analysis sample and the other zip codes are shaded by the three groups discussed above: those saving zero miles, and those above and below the median conditional on non-zero differential distance. The darker zip codes that have the largest differential distances, and are therefore closer to Community NICU or Intermediate NICU hospitals, are more likely to be in outlying areas. However, there is variation both within the major metropolitan areas and in the suburban areas with many neighboring zip codes of varying distances.

Overall, summary statistics indicate that differential distance is uncorrelated with most major observable demographic and infant health characteristics. To the extent that any differences in urban concentration are not captured by the individual controls, I examine the robustness of my results to the inclusion of zip code level controls and estimate models with zip code fixed effects in Section 6.

5 Results

5.1 OLS Estimates

OLS estimates confirm previous findings that, conditional on observable characteristics, mortality rates are on average higher for infants born in hospitals with lower level NICUs. How-

[24]Cutler (2007) and McClellan, McNeil & Newhouse (1994) also find that areas with higher differential distance to hospitals offering heart surgery are less urban.

ever, they also provide further evidence of potential unobserved differences between infants born in each type of hospital. Table 6 presents the OLS coefficient estimates of β^N, β^I, and β^C. Moving across the columns, I progressively add control variables. To account for likely similarities in health conditions and hospital choices at local levels, and because the instruments vary at the zip code level when I estimate 2SLS models, standard errors of all regression estimates in this paper are clustered by zip code. This clustering allows for arbitrary correlation of the error term within zip codes.[25]

The estimates in Column 1 reflect the unadjusted mortality differences by level of care with no additional covariates and replicate the differences in sample means from Table 3. VLBW infants born in No NICU, Intermediate NICU, and Community NICU hospitals are 7.2, 2.2, and 0.8 percentage points more likely to die, respectively, than those born in Regional NICU hospitals. The Community NICU coefficient is statistically significant at the 10% level and the other two coefficients are statistically significant at the 5% level. Column 2 adds controls for long term mortality trends in the form of year dummies and within year mortality cycles in the form of eleven month-of-year dummies and six day-of-week dummies. The estimated effect of being born in a hospital with a Community NICU increases to 1.3 percentage points and is now statistically significant at the 5% level; the other two estimates are similar to the previous column.

Column 3 adds controls for mother's demographic characteristics. The coefficient estimates decrease from Column 2 but are still positive and precisely estimated. Column 4 adds controls for the infant's baseline health characteristics and prenatal care. These covariates control for underlying health risk and are similar to controls used in previous studies. This specification estimates "risk-adjusted" mortality differences by level of care and will be treated as the baseline OLS estimates for the remainder of the paper. Except for the No NICU coefficient, the estimates in Column 4 are slightly larger than those in the previous column, and the coefficients imply that on average, infants born in hospitals with Commu-

[25]The online appendix shows that results are robust to clustering at the Hospital Service Area level and the hospital level as well.

nity NICUs, Intermediate NICUs, or No NICUs are 1.2, 2.1, or 5.0 percentage points more likely to die than those born in hospitals with Regional NICUs, respectively. Relative to the sample mean mortality rate of 15.7%, these coefficients imply effects of 7.6%, 13.4%, and 31.8%, respectively.

OLS estimates lead to the conclusion that infants born in lower level hospitals experience higher risk-adjusted mortality rates, confirming the previous literature. Infants born in No NICU hospitals have the highest risk-adjusted mortality rate, and most relevant to the trend towards deregionalization, infants born in Intermediate and Community NICU hospitals experience statistically and qualitatively higher mortality rates than those born in Regional NICU hospitals. However, the finding that the coefficient estimates are sensitive to controls implies that observed determinants of mortality are correlated with level of care. The fact that the coefficient estimates increase or decrease depending on which controls are added reinforces that the direction of any selection is ambiguous. Evidence of selection on the observables emphasizes the importance of accounting for any potential unobserved selection as well.

5.2 First Stage Estimates

This section presents estimates of the first stage relationships between distance and the level of care chosen described by Equation (2). These estimates provide evidence that the distance measures are both valid and strong instruments. Table 7 presents estimates of the first stage coefficients, building up to the baseline specification by progressively adding controls across the columns for each level of care. The coefficient estimates and standard errors show little to no change across columns. This finding implies little correlation between distance and observable characteristics and further supports the hypotheses that the instruments are uncorrelated with unobserved characteristics as well.[26]

[26] The online appendix provides two additional tests of the validity of the instruments. First, in a similar test of the correlation between the instruments and observable characteristcs, I shows that reduced form estimates of the effect of distance on neonatal mortality are also insensitive to which controls are included. Second, I show that distance has a large effect on level of care for infants in other birth weight categories,

Distances are also sgrong predictors of the level of care chosen. Columns 4, 8, and 12, present the main first stage specifications with all controls included. All of the first stage coefficient estimates are strongly statistically significant and show the expected substitution patterns. Individuals living closer to a particular level of care are more likely to choose that level of care and less likely to choose the other levels of care. For example, a ten mile increase in *ComDist*, associated with living ten miles closer to a Community NICU or higher, decreases the probability of choosing a No NICU hospital and an Intermediate NICU hospital by 2.5 and 2.7 percentage points, respectively, and increases the probability of choosing a Community NICU hospital by 7.4 percentage points.[27] These coefficient estimates are equivalent to 33%, 24%, and 31% changes relative to their respective level of care indicator sample means. These are large effects given the standard deviations of the distance instruments are around twenty. Qualitatively, distance is an important determinant of the level of care a mother chooses.

Below the estimates in each panel I report F-Statistics testing the null that the three distance coefficients are jointly equal to zero. The F-Statistics for the main specifications with the full set of controls are 32.46, 44.56, and 38.35, all well above the rule-of-thumb cutoff of 10 typically used to assess finite sample bias from weak instruments. Additionally, the fact that each instrument is significant in all three equations and has a particularly large coefficient estimate in the equation corresponding to its respective level of care, suggests that each of the three instruments provide independent variation to identify the model.

5.3 2SLS Estimates

In contrast with the OLS results, 2SLS estimates reveal little difference in mortality by level of neonatal intensive care available at the birth hospital when unobserved selection is

supporting the claim that NICUs are used to attract mothers of all risk levels.

[27]Though not a part of the estimation, there is implicitly a fourth relationship between the probability of choosing a hospital with a Regional NICU and distance. While not shown in the table, this same change decreases the probability of choosing a hospital with a Regional NICU by 2.2 percentage points, confirming the findings of Haberland, Phibbs & Baker (2006) that lower level NICUs divert high-risk births from Regional NICUs.

taken into account. Table 8 reports the 2SLS results. Column 1 repeats the baseline OLS results with all controls from Table 6. All three 2SLS coefficient estimates in Column 2 are substantially lower than their counterparts in Column 1. The coefficient of the No NICU indicator decreases from 0.050 to -0.030, the coefficient of the Intermediate NICU indicator decreases from 0.021 to 0.009, and the coefficient of the Community NICU indicator decreases from 0.012 to -0.063. The No NICU and Community NICU coefficient estimates actually change signs and the Intermediate NICU coefficient estimate falls by half, but the standard errors increase by a factor of between three and nine. The Community NICU coefficient estimate is marginally statistically significant (at the 10% level), but neither of the other two estimates in Column 2 are statistically significant.

Despite the large standard errors, the 2SLS estimates are clearly different from and bounded below the OLS estimates. First, I conduct a Hausman test of the null hypothesis that both the OLS and 2SLS estimates are consistent against the alternative that only the 2SLS estimates are consistent.[28] The p-value of this test is 0.031, so the null is rejected at the 5% significance level. This test implies that the 2SLS coefficient estimates are statistically different from the OLS estimates and provide more consistent estimates.

Second, even the upper bounds of the 2SLS estimates imply much lower quantitative and qualitative effects on mortality than the OLS estimates, particularly for the No NICU and Community NICU coefficients. Figure 1 plots the OLS and 2SLS coefficient estimates scaled by mean neonatal mortality. It also plots one and two standard deviation intervals above the 2SLS coefficient estimates. The OLS coefficient estimate of the No NICU coefficient implies 31.8% higher mortality relative to being born in a Regional NICU hospital. The 2SLS coefficient estimate is large and negative, one standard deviation above the 2SLS coefficient

[28]The usual Hausman test also assumes that the OLS estimates are efficient under the null hypothesis. However, clustered standard errors result in a covariance matrix that is not asymptotically efficient. Therefore, I construct the Hausman test statistic using a paired bootstrap strategy that samples at the zip code level. The sample has 1,144 zip codes, so I construct 5,000 random samples of the data that each draw 1,144 zip codes with replacement. For each bootstrap sample, I run my OLS and 2SLS regressions and save the coefficient estimates. I then construct the estimated variance-covariance matrix of the difference between the OLS and 2SLS coefficients based on the distribution of these 5,000 estimates. See A. Colin Cameron & Pravin K. Trivedi (2005, p. 378) for details.

estimate is still below zero, and even two standard deviations above implies an effect of 17.9% – 44% lower than the OLS estimate. Similarly, one standard deviation above the Community NICU coefficient estimate is still far below zero, and two standard deviations above implies an effect of 4% − 46% lower than the OLS effect of 7.4%. One standard deviation above the Intermediate NICU 2SLS coefficient estimate is above the OLS estimate, but the point estimate is still 55% lower than the OLS point estimate.

The 2SLS estimates are not statistically different from zero and are small in magnitude compared to OLS estimates. This finding provides evidence that the OLS estimates of higher mortality at the three lower levels of care relative to Regional NICU hospitals are overstated. The dominant form of selection is unobservably higher risk births occurring in lower level hospitals. These results imply that policy measures aimed at reversing the effects of deregionalization are likely to have a limited impact on mortality. Relocating mothers who would have chosen to give birth in lower level hospitals to Regional NICU hospitals prior to birth would not improve mortality rates because the relocated deliveries would be from the less healthy portion of the distribution.

It is important to emphasize that I am estimating how the level of care at the hospital in which an infant is *born* impacts mortality. My results do not imply that being *treated* in a hospital with a higher level NICU has no effect on outcomes. In fact, a likely mechanism behind my results is that infants born in hospitals with lower levels of care achieve similar outcomes to those born in higher level hospitals because the former group will be transferred to a higher level hospital after birth if necessary. In my sample 66% of VLBW infants born in hospitals with No NICUs or Intermediate NICUs are transferred to a higher level hospital after birth, and 85% of those that are transferred are sent to a Regional NICU hospital.

In order to explore whether the probability of transfer is systematically impacted by distance, I regress an indicator for whether or not an infant is transferred to a Regional NICU hospital on the three distance instruments for the sample of VLBW infants born in No NICU or Intermediate NICU hospitals. Statistically insignificant coefficients on the three distance instruments would suggest transfers occur when medically necessary and are not impacted by where a mother lives in relation to where NICUs are located. Results of this regression do reveal a positive and statistically significant coefficient of 0.029 on the No NICU distance variable, but the coefficient estimates on the other two distance instruments are very small and statistically insignificant (-0.00008 and 0.008, respectively). When I instead regress the transfer indicator on all four distance measures instead of the three differential distance measures, I find that the positive coefficient on the No NICU differential distance variable is being driven by lower transfer probabilities for mothers living close to Regional NICUs. This finding is likely a result of using the selected sample of infants born in No NICU or Intermediate NICU hospitals. Infants of mothers who live close to Regional NICU hospitals, but choose not to deliver in the Regional NICU hospital are likely to have healthier infants and less medical need for transfer. Overall, these results suggest that I find no gradient between level of care at the birth hospital and mortality because VLBW infants are transferred to hospitals with higher levels of care when medically necessary, and the location of lower level NICUs does not change the probability of eventually being treated in a hospital with a higher level NICU.[29]

6 Robustness Tests

6.1 Zip Code of Residence Controls and Fixed Effects

Sample means by differential distance in Table 5 showed that individuals living closer to each of the three lower levels of care relative to Regional NICU hospitals live in zip codes with lower population density and higher income. Though I control for many individual level covariates in my main results, if these zip code level characteristics are conditionally

[29] Recall that my analysis sample only includes mothers residing in urban and suburban zip codes. These results do not address the relationship between access to care and health outcomes for mothers residing in rural areas. Understanding how transfer networks function for rural residents is an interesting topic for future research.

correlated with distance and infant health, 2SLS estimates would be biased. Therefore, I test the robustness of my estimates to controlling for zip code level population density; percent black; percent Hispanic; percent of the population over 25 with no college, some college, a college degree, and more than a college degree; and median household income.[30]

Additionally, distances may factor into the hospital choice decision differently in urban and suburban areas. For example, five miles in downtown Los Angeles may have a much different travel time than five miles in a suburban area. Furthermore, hospitals are located closer to each other in more urban areas than less urban areas. Using differential distance and controlling for all three distance variables captures some of these features, but the fist stage regression may have more predictive power if the effect of distance is allowed to vary with population density. I therefore estimate models with interactions of the distance measures and zip code population density added to the instrument set.

Panel B of Table 9 presents the OLS and 2SLS squares results, with each row listing coefficient estimates from one regression.[31] The baseline estimates are repeated in Panel A. The OLS and 2SLS coefficient and standard error estimates in specifications controlling for zip code level characteristics are very similar to the baseline estimates. Controlling for differences between urban and suburban zip codes does not impact the results. The last row of Panel B presents results when the instrument set includes interactions with population density. The standard errors of these estimates are very similar to the specification with zip code level controls and the baseline specification; however, the coefficients all move towards zero and none are statistically significant. If anything, allowing the effect of distance to differ with population density results in point estimates that are even closer to zero.

Next, I estimate models with zip code of residence fixed effects to control for any other characteristics that are constant within a zip code, but not accounted for by the census data controls. Identification with these fixed effects comes from changes over time in a zip code's

[30]All zip code level variables are calculated from the 2000 Census. Unfortunately, the 1990 census does not provide comparable data at the zip code level.

[31]The online appendix shows that the instruments are still strong predictors of hospital choice with these additional control variables and with distance interacted with population density.

distances to each level of care caused by new, upgraded, or closed NICUs nearby during the sample period. Thus, the variation in distance is directly driven by deregionalization during the sample period. 25% of the VLBW sample lives in a zip code that at some point between 1991 and 2001 experiences a change in Intermediate Distance, and the average change is 4.5 miles. 32% lives in a zip code that experiences a change in Community Distance, and the average change is 3.9 miles.

With fixed effects the instruments are valid if zip code level changes in distance are uncorrelated with zip code level changes in unobserved mortality.[32] Even if zip codes at different distances differ systematically, identification will only be threatened if unobserved mortality trends are conditionally correlated with changes in distance. The online appendix shows shows that at least trends in mean observable demographic and underlying health variables do not systematically differ between zip codes experiencing different changes in distance. This finding of parallel trends is not surprising given the evidence that deregionalization has not been driven by the health needs of high-risk infants.

Panel C of Tables 9 shows the results.[33] OLS results are similar to the cross sectional results. The 2SLS estimates are again not statistically significant. The fixed effects lead to much larger standard errors and more negative point estimates of the No NICU and Community NICU coefficients, but the qualitative results are similar: negative or small point estimates, indicating no difference in mortality outcomes by level of care at the birth hospital. When the instruments are allowed to vary with population density, the negative point estimates of the No NICU and Community NICU coefficients are cut by about two thirds and move towards zero as in the specifications without fixed effects. These specifications confirm that the main results are robust to the most complete possible controls for local characteristics. They also show that the cross sectional 2SLS specifications estimate similar effects to specifications identified directly from changes in distance related to deregionalization.

[32]Formally, the assumption is $E[\ddot{D}'_{zt}\ddot{\varepsilon}_{izt}|\ddot{\mathbf{X}}_{izt}] = \mathbf{0}$, where the dots indicate variables in deviation-from-zip-code-mean-form.

[33]The online appendix presents first stage results with zip code of residence fixed effects.

6.2 Pooling No NICUs and Intermediate NICUs

I also estimate models where I pool No NICU and Intermediate NICU hospitals into one category. Only 7.6% and 11.1% of the VLBW sample are born in these two types of hospitals, respectively. Thus combining them into one group may provide more precision. Additionally, some of the first stage predictions of these indicators are outside the unit interval. Pooling these two groups reduces the percentage of observations with at least one of their first stage predictions outside the unit interval from 12.8% to 2.7%.[34] In Panel D of Table 9 OLS estimates are as expected, with a similar Community NICU coefficient estimate to the baseline specification and coefficient estimates of the pooled No/Intermediate NICU coefficient between the original No NICU and Intermediate NICU coefficient estimates.[35] The precision gains in the 2SLS estimates are not large, but the point estimates are closer to zero, and none of them are statistically significant negative estimates.

6.3 Heterogeneity and Local Average Treatment Effects

Throughout the paper I have assumed a homogeneous effect of level of care on mortality for all VLBW infants. However, it is possible that the effect may vary by the infant's characteristics. This is particularly important with instrumental variable estimates because they only estimate the impact of level of care on mortality for the sub-group of infants whose mothers choose level of care based on the instruments. If these "compliers" are different from the rest of the sample, these estimates will represent a local average treatment effect (LATE) (Joshua D. Angrist, Guido W. Imbens & Donald B. Rubin 1996, Guido W. Imbens & Joshua D. Angrist 1994). One cannot directly observe the compliers in the data, but I estimate my OLS and 2SLS regression equations on various sub-samples based on observable characteristics to ensure the estimates are not being driven by any particular

[34]It is also likely medically reasonable to pool these two groups. Neither of these types of hospitals is designed to care for VLBW infants, and neither provides mechanical ventilation. Additionally, about 60% of VLBW infants born at these two levels are transferred to Regional NICUs. In contrast, only 20% of infants born in Community NICU hospitals are transferred to Regional NICUs.

[35]The online appendix presents first stage results.

groups. Understanding any heterogeneity is also important for policy implications. If there is a gradient between level of care and mortality for some groups, interventions may be warranted to target these specific groups and ensure they are able to deliver in higher level hospitals.

Panel A of Table 10 presents results for various subsamples with the baseline estimation from Table 8 repeated in Column 1. Overall, the OLS and 2SLS coefficient estimates are similar across all reported sub-groups. Column 2 shows the results for infants of Hispanic mothers. The 2SLS estimate of the effect of being born in an Intermediate NICU hospital (0.025) is close to the OLS coefficient estimate (0.028) for this group, but still statistically insignificant. The other two 2SLS coefficient estimates are negative, statistically insignificant, and similar to the baseline sample.

Column 3 excludes infants of black mothers from the estimation. Infants with black mothers make up a small subset of the sample, so I do not estimate the regressions for them alone, but excluding them does not have much effect on the estimates. This sample also provides a useful robustness check because of the difference in percent black by differential distance reported in Table 5. The estimates for the population of infants with mothers covered by Medicaid in Column 4 are similar to the baseline specification, indicating the results are similar by insurance coverage. The sample of infants whose mothers have no college education in Column 5 has a 2SLS Intermediate NICU coefficient that is the same as the OLS estimate (0.027), but again it is not statistically significant and the other 2SLS coefficient estimates are negative.

In the previous section I show the results are robust to controlling for population density and allowing the effect of distance to differ by population density. One might also be concerned that the results are driven by either urban or suburban areas, which I address in Column 6. This column presents estimates for the sub-sample whose zip code population density is below the median. Again, the 2SLS coefficient on being born in an Intermediate NICU hospital (0.022) is similar to the OLS coefficient (0.028), but the other estimates are similar to the baseline specification, indicating the results are similar for individuals in urban and suburban areas.

Finally, the effect of level of care may have changed over time. Mortality rates for VLBW infants decreased during the early 1990s, but leveled off during the latter part of the decade (Jeffrey D. Horbar, Gary J. Badger, Joseph H. Carpenter, Avroy A. Fanaroff, Sarah Kilpatrick, Meena LaCorte, Roderic Phibbs & Roger F. Soll 2002).[36] Also, Table 2 shows that the diffusion of NICUs leveled off during the second half of the decade. It is possible that the gradient between level of care and mortality changed during this time period if technology improved, if new NICUs improved over time due to learning, or if the propensity for lower level units to transfer infants to higher levels changed over time. Column 7 presents results for births occurring during the first half of the sample from 1991 to 1995. The OLS gradient between level of care is greater during this time period as compared to results for the full time period, but because mean mortality was higher during the earlier period, the relative effects are similar. The 2SLS estimates are similarly small and statistically insignificant as compared to the baseline estimation. There is no evidence of a differential effect of level of care on mortality over time.

Despite evidence that the effect of level of care on mortality does not vary by observable characteristics, there still may be unobserved heterogeneity, leading to estimation of a LATE for an unidentifiable group of compliers. That being said, because the compliers are the infants whose mothers choose their delivery hospital based on distance, the LATE would in fact be the policy relevant effect. Even if the 2SLS estimates do not represent the effect of level of care on mortality for the entire population, they still imply that the population that would be impacted by policy measures regarding the geographic distribution of NICUs does not experience different mortality rates by level of care.

[36] In my sample, mean neonatal mortality fell from 20.08% to 14.80% between 1991 and 1995, but only fell to 13.62% by 2001.

7 Conclusion

This paper estimates the causal effects of level of neonatal intensive care at the birth hospital on VLBW mortality. Evidence of higher risk-adjusted mortality rates for VLBW infants born in hospitals with lower level NICUs has led advocates to suggest high-risk mothers be referred to more sophisticated hospitals prior to delivery. However, these estimates could be biased in either direction by unobserved selection. To overcome selection concerns, I utilize an instrumental variables strategy that exploits exogenous variation in distance from a mother's residence to the nearest hospital of each level of care. NICU location has been driven by factors unrelated to the health of VLBW infants, and I provide evidence in my data that distance is uncorrelated with health conditions.

My OLS estimates confirm the previous literature and imply 7.6%, 13.4%, and 31.8% higher risk-adjusted mortality rates for VLBW infants born in Community, Intermediate, and No NICU hospitals, respectively, relative those born in Regional NICUs hospitals. However, my instrumental variables estimates imply that these mortality differences are overstated. 2SLS estimates are not statistically different from zero and are small in magnitude. The No NICU and Community NICU 2SLS coefficient estimates are bounded well below their OLS counterparts, with even two standard deviations above the 2SLS estimates lying about 50% below the OLS estimates. The Intermediate NICU 2SLS coefficient estimate is not clearly bounded below the OLS estimate, but the point estimate is half the magnitude. My results are robust to controlling for zip code level characteristics and zip code level fixed effects. I also find no evidence that the effect of level of care on mortality is heterogeneous by demographics. Even if the effect varies on other unobservable dimensions, any unobserved heterogeneity would lead to a local average treatment effect directly identified from infants impacted by deregionlization.

The finding that the 2SLS estimates are below the OLS estimates, reveals that mothers with higher unobserved risk select into hospitals with lower levels of care. In terms of mortality, these results imply that relocating high-risk deliveries to Regional NICU hospitals prior to birth will not result in improved health outcomes. Instead, Regional hospitals would be treating new patients with higher unobserved acuity. I also show evidence that level of care at the birth hospital does not impact mortality because infants born in No NICU and Intermediate NICU hospitals are often transferred to Regional NICU hospitals, and these transfers are independent of how close mothers live to lower level facilities. Deregionalization does not appear to prevent infants born in No NICU or Intermediate NICU hospitals from eventually receiving care in Regional NICUs.

This analysis has addressed the first-order question of how deregionalization has impacted VLBW mortality. Future research is needed to understand the full welfare impacts of this trend. First, while mortality may not vary by level of care at the birth hospital, there may be important differences in cost of care. If larger hospitals achieve economies of scale, they may be more efficient in treating sick infants. Inter-hospital transfers may also be costly, both monetarily and emotionally. Alternatively, more sophisticated facilities may provide more costly procedures with little marginal return. Second, there may be important effects of deregionalization on quality and cost of care for healthier infants.

Third, if mothers value shorter travel time and more convenient visitation of family members, access to at least some level of intensive care at nearby hospitals may increase utility. Also, more competition in the neonatal intensive care market may lead to lower prices. Vivian Ho, Robert J. Town & Martin J. Heslin (2007) study the market for Whipple surgery, a treatment for pancreatic cancer, and find that regionalizing this treatment by consolidating it to the hospitals with the highest volume leads to substantial price increases.[37] Finally, further research is warranted to understand the determinants of NICU adoption by hospitals and whether hospitals are able to recoup their fixed costs by attracting profitable patients.

[37]These authors do find that regionalization of Whipple surgery can reduce mortality, but price increases cancel out over half of the increased consumer surplus.

References

Almond, Douglas, and Joseph J. Doyle. 2008. "After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays." *NBER Working Paper*, No. 13877.

Almond, Douglas, Joseph J. Doyle, Amanda E. Kowalski, and Heidi Williams. 2008. "Estimating Marginal Returns to Medical Care: Evidence from At-Risk Newborns." *NBER Working Paper*, No. 14522.

Almond, Douglas, Kenneth Y. Chay, and David S. Lee. 2005. "The Costs of Low Birth Weight." *Quarterly Journal of Economics*, 120(3): 1031–1083.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113(1): 151–184.

Angrist, Joshua D. 2001. "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice." *Journal of Business & Economic Statistics*, 19(1): 2–16.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*, 91(434): 444–455.

Baicker, Katherine, and Amitabh Chandra. 2004. "Medicare Spending, The Physician Workforce, And Beneficiaries' Quality Of Care." *Health Affairs*, W4–184–W4–197.

Baicker, Katherine, Kasey S. Buckles, and Amitabh Chandra. 2006. "Geographic Variation In The Appropriate Use Of Cesarean Delivery." *Health Affairs*, 25(5): w355–367.

Baker, Laurence C., and Ciaran S. Phibbs. 2002. “Managed Care, Technology Adoption, and Health Care: The Adoption of Neonatal Intensive Care.” *The RAND Journal of Economics*, 33(3): 524–548.

Behrman, Richard E., and Adrienne Stith Butler, ed. 2007. *Preterm Birth: Causes, Consequences, and Prevention.* Washington, D.C.:The National Academies Press.

Bhattacharya, Jay, Dana Goldman, and Daniel McCaffrey. 2006. “Estimating Probit Models with Self-Selected Treatments.” *Statistics in Medicine*, 25(3): 389–413.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications.* . 1st ed., New York:Cambridge University Press.

Chandra, Amitabh, and Douglas O. Staiger. 2007. “Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks.” *Journal of Political Economy*, 115: 103–140.

Cifuentes, Javier, Janet Bronstein, Ciaran S. Phibbs, Roderic H. Phibbs, Susan K. Schmitt, and Waldemar A. Carlo. 2002. “Mortality in Low Birth Weight Infants According to Level of Neonatal Care at Hospital of Birth.” *Pediatrics*, 109(5): 745–751.

Committee on Fetus and Newborn. 2004. “Levels of Neonatal Care.” *Pediatrics*, 114(5): 1341–1347.

Committee on Perinatal Health. 1976. *Toward Improving the Outcome of Pregnancy: Recommendations for the Regional Development of Maternal and Perinatal Health.* White Plains, NY:March of Dimes.

Committee on Perinatal Health. 1993. *Toward Improving the Outcome of Pregnancy: The 90s and Beyond.* White Plains, NY:March of Dimes.

Cutler, David M. 2007. "The Lifetime Costs and Benefits of Medical Technology." *Journal of Health Economics*, 26(6): 1081–1100.

Cutler, David M., and Ellen Meara. 2000. "The Technology of Birth: Is It Worth It?" In *Frontiers in Health Policy Research, Volume 3.* , ed. Alan M. Garber, 33–68. National Bureau of Economic Research, Inc.

Cutler, David M., and Mark McClellan. 2001. "Productivity Change in Health Care." *The American Economic Review*, 91(2): 281–286.

Dafny, Leemore S. 2005. "Games Hospitals Play: Entry Deterrence in Hospital Procedure Markets." *Journal of Economics & Management Strategy*, 14(3): 513–542.

Dranove, David, Mark Shanley, and Carol Simon. 1992. "Is Hospital Competition Wasteful?" *The RAND Journal of Economics*, 23(2): 247–262.

Evans, William N., and Craig L. Garthwaite. 2009. "Estimating Heterogeneity in the Benefits of Medical Treatment Intensity." *National Bureau of Economic Research Working Paper Series*, No. 15309.

Evans, William N., Craig Garthwaite, and Heng Wei. 2008. "The Impact of Early Discharge Laws on the Health of Newborns." *Journal of Health Economics*, 27(4): 843–870.

Fisher, Elliott S., David E. Wennberg, Therese A. Stukel, Daniel J. Gottlieb, F. L. Lucas, and Etoile L. Pinder. 2003*a*. "The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care." *Annals of Internal Medicine*, 138(4): 273–287.

Fisher, Elliott S., David E. Wennberg, Therese A. Stukel, Daniel J. Gottlieb, F. L. Lucas, and Etoile L. Pinder. 2003*b*. "The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care." *Annals of Internal Medicine*, 138(4): 288–298.

Friedman, Bernard, Kelly J. Devers, Claudia A. Steiner, and Steven H. Fox. 2002. "The Use of Expensive Health Technologies in the Era of Managed Care: The Remarkable Case of Neonatal Intensive Care." *Journal of Health Politics, Policy and Law*, 27(3): 441–464.

Fuchs, Victor R. 2004. "Perspective: More Variation In Use Of Care, More Flat-Of-The-Curve Medicine." *Health Affairs*, VAR104–VAR107.

Gaynor, Martin. 2006. "What Do We Know About Competition and Quality in Health Care Markets?" *NBER Working Paper*, No. 12301.

Goldman, Dana, and John A. Romley. 2008. "Hospitals As Hotels: The Role of Patient Amenities in Hospital Demand." *NBER Working Paper*, No. 14619.

Goodman, David C., Elliott S. Fisher, George A. Little, Therese A. Stukel, and Chiang hua Chang. 2001. "Are Neonatal Intensive Care Resources Located According to Need? Regional Variation in Neonatologists, Beds, and Low Birth Weight Newborns." *Pediatrics*, 108(2): 426–431.

Gowrisankaran, Gautam, and Robert J. Town. 2003. "Competition, Payers, and Hospital Quality." *Health Services Research*, 38(6p1): 1403–1422.

Gowrisankaran, Gautam, Vivian Ho, and Robert J. Town. 2006. "Causality, Learning and Forgetting in Surgery." *Working Paper*.

Haberland, Corinna A., Ciaran S. Phibbs, and Laurence C. Baker. 2006. "Effect of Opening Midlevel Neonatal Intensive Care Units on the Location of Low Birth Weight Births in California." *Pediatrics*, 118(6): e1667–1679.

Hall, Robert E., and Charles I. Jones. 2007. "The Value of Life and the Rise in Health Spending." *Quarterly Journal of Economics*, 122(1): 39–72.

Horbar, Jeffrey D., Gary J. Badger, Joseph H. Carpenter, Avroy A. Fanaroff, Sarah Kilpatrick, Meena LaCorte, Roderic Phibbs, and Roger F. Soll. 2002. "Trends in Mortality and Morbidity for Very Low Birth Weight Infants, 1991-1999." *Pediatrics*, 110(1): 143–151.

Horwitz, Jill R. 2005. "Making Profits And Providing Care: Comparing Nonprofit, For-Profit, And Government Hospitals." *Health Affairs*, 24(3): 790–801.

Ho, Vivian, Robert J. Town, and Martin J. Heslin. 2007. "Regionalization Versus Competition in Complex Cancer Surgery." *Health Economics, Policy and Law*, 2(01): 51–71.

Howell, Embry M., Douglas Richardson, Paul Ginsburg, and Barbara Foot. 2002. "Deregionalization of Neonatal Intensive Care in Urban Areas." *American Journal of Public Health*, 92(1): 119–124.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, 62(2): 467–475.

Kessler, Daniel P., and Mark B. McClellan. 2000. "Is Hospital Competition Socially Wasteful?" *Quarterly Journal of Economics*, 115(2): 577–615.

Luce, Bryan R., Josephine Mauskopf, Frank A. Sloan, Jan Ostermann, and L. Clark Paramore. 2006. "The Return on Investment in Health Care: From 1980 to 2000." *Value in Health*, 9(3): 146–156.

McClellan, Mark, and Joseph P. Newhouse. 1997. "The Marginal Cost-Effectiveness of Medical Technology: A Panel Instrumental-Variables Approach." *Journal of Econometrics*, 77(1): 39–64.

McClellan, Mark, Barbara J. McNeil, and Joseph P. Newhouse. 1994. "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality? Analysis Using Instrumental Variables." *Journal of the American Medical Association*, 272(11): 859–866.

McCormick, Marie C., and Douglas K. Richardson. 1995. "Access to Neonatal Intensive Care." *The Future of Children*, 5(1): 162–175.

Murphy, Kevin M., and Robert H. Topel. 2003. "The Economic Value of Medical Research." In *Measuring the Gains from Medical Research: An Economic Approach.* , ed. Kevin M. Murphy and Robert H. Topel. Chicago:University of Chicago Press.

Phibbs, Ciaran S., David H. Mark, Harold S. Luft, Deborah J. Peltzman-Rennie, Deborah W. Garnick, Erik Lichtenberg, and Stephen J. McPhee. 1993. "Choice of Hospital for Delivery: A Comparison of High-Risk and Low-Risk Women." *Health Services Research*, 28(2): 201222.

Phibbs, Ciaran S., Janet M. Bronstein, Eric Buxton, and Roderic H. Phibbs. 1996. "The Effects of Patient Volume and Level of Care at the Hospital of Birth on Neonatal Mortality." *Journal of the American Medical Association*, 276(13): 1054–1059.

Phibbs, Ciaran S., Laurence C. Baker, Aaron B. Caughey, Beate Danielsen, Susan K. Schmitt, and Roderic H. Phibbs. 2007. "Level and Volume of Neonatal Intensive Care and Mortality in Very-Low-Birth-Weight Infants." *New England Journal of Medicine*, 356(21): 2165–2175.

Russell, Rebecca B., Nancy S. Green, Claudia A. Steiner, Susan Meikle, Jennifer L. Howse, Karalee Poschman, Todd Dias, Lisa Potetz, Michael J. Davidoff, Karla Damus, and Joann R. Petrini. 2007. "Cost of Hospitalization for Preterm and Low Birth Weight Infants in the United States." *Pediatrics*, 120(1): e1–9.

Schmidt-Dengler, Philipp. 2006. "The Timing of New Technology Adoption: The Case of MRI." *Working Paper*.

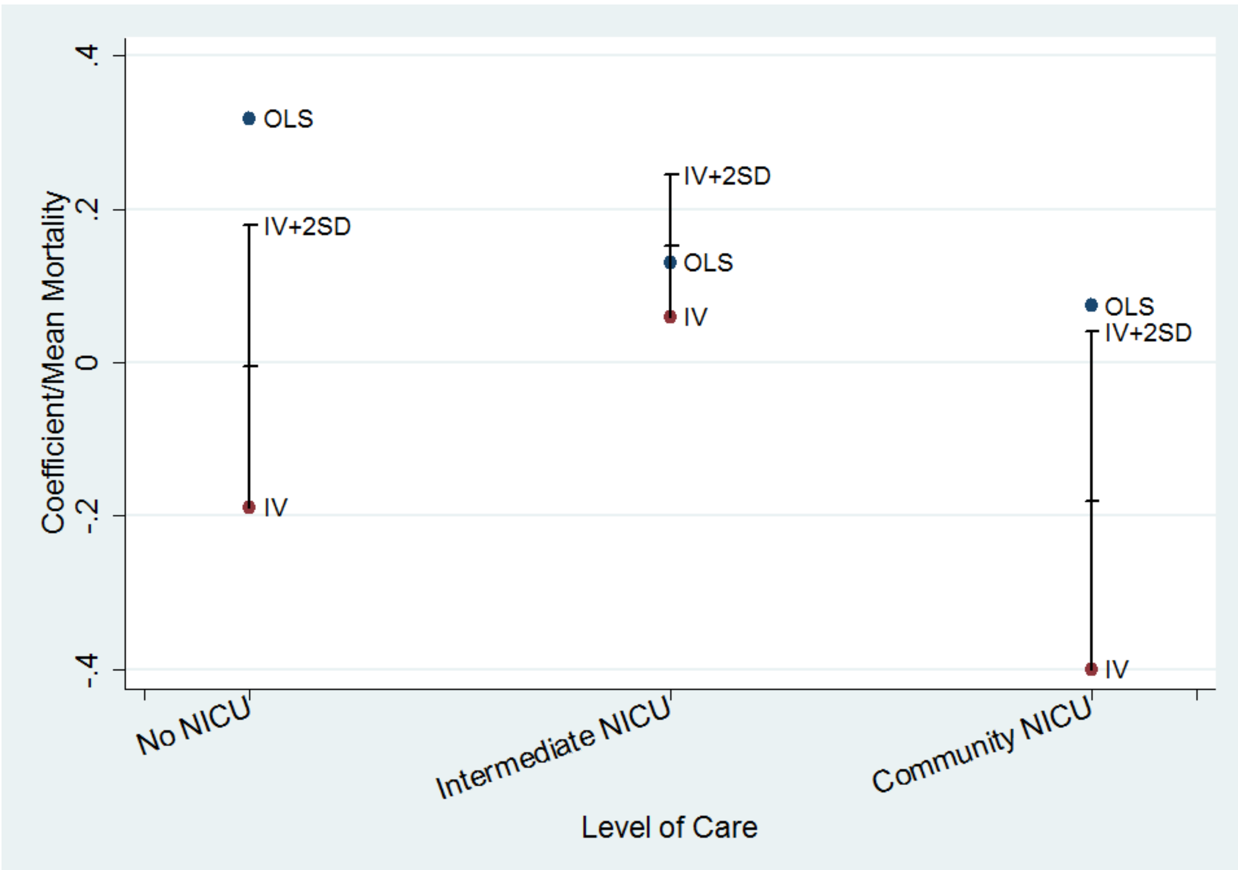
Schwartz, Rachel M. 1996. "Supply and Demand for Neonatal Intensive Care: Trends and Implications." *Journal of Perinatology*, 16(6): 483–489.

Schwartz, Rachel M., Russell Kellogg, and Janet H. Muri. 2000. "Specialty Newborn Care: Trends and Issues." *Journal of Perinatology*, 20(8 Pt 1): 520–529.

Tay, Abigail. 2003. "Assessing Competition in Hospital Care Markets: The Importance of Accounting for Quality Differentiation." *The RAND Journal of Economics*, 34(4): 786–814.

Wooldridge, Jeffrey M. 2001. *Econometric Analysis of Cross Section and Panel Data.* . 1st ed., Cambridge, MA:The MIT Press.

Figure 1: Coefficient Estimate Magnitudes



Notes: This figure plots the OLS and 2SLS coefficient estimates from Table 8 divided by mean neonatal mortality (15.7%). The dashed points indicate one and two standard deviation intervals above the 2SLS coefficient estimates.

Table 1: Detailed Level of Care Definitions

Level	Care Provided
I	Basic neonatal care for healthy infants
	No Intensive Care Unit
II	Have an intensive care unit
	Care for midly ill infants
	Do not provide mechanical ventilation
IIIA	Provide mechanical ventilation with restrictions
	(e.g., only for less than 96 hours, or only for infants
	weighing above 1,000 grams)
IIIB	Provide mechanical ventilation without restrictions
IIIC	Provide major neonatal surgery excluding cardiac surgery requiring
	bypass and/or extracorporeal membrane oxygenation (ECMO)
IIID	Provide cardiac surgery requiring bypass and/or ECMO

Notes: Level of neonatal care definitions from Phibbs et al. (2007). There are three ICD-9 CM codes indicating mechanical ventilation: <96 hours, >96 hours, and duration unknown. Hospitals with NICU beds that do not have occurrences of any of these codes are labeled as Level II. In distinguishing between Level IIIA and IIIB, Phibbs et al. (2007) count units that never provide ventilation for more than 96 hours as IIIA. For units that provide both types but do not provide any surgery, they examine the patterns of ventilation by duration and birth weight to distinguish which appear to have restrictions.

Table 2: California Obstetric Hospitals by Year and Level of Care

Year	No NICU	Intermediate NICU	Community NICU	Regional NICU	Total
1991	161	58	35	42	296
1992	153	52	43	44	292
1993	149	53	45	45	292
1994	147	56	45	45	293
1995	148	49	51	46	294
1996	140	48	54	46	288
1997	141	47	55	46	289
1998	139	45	58	46	288
1999	135	44	60	46	285
2000	130	45	57	45	277
2001	122	45	57	45	269

Notes: Author's tabulations based on data from Phibbs et al. (2007) and OSHPD Annual Utilization Files. See level of care definitions in text.

Table 3: Sample Means by Level of Care at Birth Hospital

	No NICU	Intermediate NICU	Community NICU	Regional NICU
Mortality				
28 Day Mortality	0.202	0.150	0.139	0.131
1 Year Mortality	0.235	0.185	0.167	0.160
Neonatal Mortality	0.219	0.169	0.155	0.147
Mother's Demographics				
Age	25.781	26.861	27.939	28.084
Black	0.098	0.205	0.128	0.186
Hispanic	0.567	0.374	0.454	0.434
Medicaid	0.591	0.546	0.455	0.508
HMO	0.148	0.213	0.276	0.212
Self Pay	0.095	0.045	0.031	0.022
No College	0.788	0.679	0.643	0.654
Some College	0.151	0.199	0.195	0.183
College	0.061	0.122	0.162	0.163
Infant Characteristics				
Month Prenatal Care Began	2.323	2.321	2.190	2.202
# of Prenatal Visits	6.692	8.209	8.718	8.873
Parity	2.349	2.358	2.209	2.289
Male	0.542	0.521	0.512	0.511
Multiple Birth	0.167	0.210	0.218	0.244
Birth Weight (Grams)	1067.017	1063.166	1064.203	1055.371
Gestation (Weeks)	30.079	30.083	29.836	29.928
Clinical Condition	0.153	0.192	0.237	0.306
Small for Gest.	0.034	0.049	0.065	0.055
Large for Gest.	0.008	0.009	0.012	0.021
Treatment				
Total Length of Stay	39.179	44.197	50.828	53.319
Total Charges ($1,000s)	156.595	158.987	204.456	228.216
Charges/Day ($1,000s)	1.656	2.894	4.059	4.136
Ventilation	0.136	0.235	0.571	0.556
Transfer	0.706	0.638	0.209	0.114
Observations	3,268	4,788	10,136	24,720
# of Hospitals	142	49	51	45

Notes: Columns display sample means for infants delivered in hospitals at four levels of care. Total Length of Stay and Total Charges sum length of stay and hospital charges over all contiguous hospitalizations prior to first being discharged home or dying. Neonatal mortality is mortality within twenty-eight days of birth or within one year if an infant is continuously hospitalized since birth. Number of hospitals indicates the average number of hospitals providing each level of care over the 11-year sample. See Table 2 for number of hospitals by year.

Table 4: Summary Statistics of Distance Variables

	Mean	SD
D(No+)	3.673	4.206
D(Int+)	5.709	8.065
D(Com+)	8.064	11.983
D(Reg)	14.830	22.991
NoDist	11.156	21.723
IntDist	9.120	20.249
ComDist	6.766	18.446
N	42,912	

Notes: The first four rows show the mean and standard deviation of distance to the nearest hospital offering each level of care or higher. The next three rows show the mean and standard deviation of differential distance to the nearest hospital offering each level of care or higher relative to the nearest Regional NICU.

Table 5: Sample Means by Distance

	By Miles Saved to Nearest No +			By Miles Saved to Nearest Int +			By Miles Saved to Nearest Com +		
	0	<Median	>Median	0	<Median	>Median	0	<Median	>Median
Distance									
Miles Saved No+	0.000	1.982	26.378	1.605	3.137	28.981	4.010	4.500	32.040
Miles Saved Int+	0.000	1.488	21.695	0.000	2.000	25.604	2.177	3.108	28.958
Miles Saved Com+	0.000	0.919	16.277	0.000	1.257	19.219	0.000	1.584	25.420
Mother's Demographics									
Age	27.826	27.913	27.517	27.787	27.936	27.490	27.752	28.038	27.411
Black	0.227	0.196	0.107	0.208	0.194	0.102	0.207	0.163	0.095
Hispanic	0.434	0.457	0.431	0.475	0.414	0.436	0.438	0.431	0.462
Medicaid	0.519	0.517	0.488	0.524	0.501	0.493	0.521	0.483	0.499
HMO	0.213	0.223	0.227	0.205	0.239	0.224	0.211	0.242	0.224
Self Pay	0.031	0.030	0.035	0.034	0.030	0.033	0.033	0.034	0.030
No College	0.684	0.668	0.650	0.694	0.648	0.650	0.676	0.645	0.661
Some College	0.170	0.181	0.198	0.171	0.189	0.197	0.178	0.188	0.197
College	0.146	0.151	0.152	0.135	0.163	0.153	0.147	0.166	0.142
Infant Characteristics									
Mth Prenatal Care Began	2.240	2.214	2.220	2.241	2.213	2.210	2.244	2.203	2.194
# of Prenatal Visits	8.196	8.733	8.679	8.277	8.792	8.728	8.537	8.628	8.683
Parity	2.390	2.247	2.259	2.370	2.229	2.246	2.316	2.228	2.270
Male	0.502	0.514	0.522	0.506	0.515	0.522	0.512	0.511	0.523
Multiple Birth	0.223	0.224	0.236	0.218	0.231	0.236	0.227	0.233	0.226
Birth Weight (Grams)	1056.087	1059.704	1060.439	1056.993	1061.276	1059.441	1060.140	1060.478	1056.107
Gestation (Weeks)	30.033	29.909	29.905	29.980	29.963	29.859	29.975	29.925	29.860
Clinical Condition	0.281	0.279	0.242	0.273	0.286	0.236	0.287	0.276	0.212
Small for Gest.	0.061	0.055	0.051	0.058	0.057	0.050	0.057	0.057	0.050
Large for Gest.	0.018	0.019	0.013	0.018	0.018	0.013	0.018	0.016	0.013
Zip Code Characteristics									
Med HH Income ($1,000)	40.511	43.736	46.599	40.221	45.705	46.690	43.265	45.441	44.694
Percent Urban	0.986	0.987	0.924	0.978	0.984	0.925	0.973	0.980	0.923
Population Density	8683.053	9752.449	3162.460	9185.203	8225.657	3312.801	8142.096	7983.427	3456.633
Observations	9,247	16,776	16,889	14,585	14,147	14,180	21,440	10,719	10,753

Notes: The first three columns display sample means by differential distance to the nearest hospital with any obstetric services, the second three columns by differential distance to the nearest Intermediate NICU or higher, and the final three columns by differential distance to the nearest Community NICU or higher. For each set of columns, the sample is divided into three groups: those with zero differential distance, and those above and below the median conditional on non-zero differential distance.

Table 6: Neonatal Mortality by Level of Care, OLS Estimates

	Dependent Variable: Neonatal Mortality			
	(1)	(2)	(3)	(4)
I(No NICU)	0.072**	0.072**	0.054**	0.050**
	(0.009)	(0.008)	(0.009)	(0.007)
I(Intermediate NICU)	0.022**	0.021**	0.017**	0.021**
	(0.007)	(0.006)	(0.006)	(0.005)
I(Community NICU)	0.008*	0.013**	0.010**	0.012**
	(0.004)	(0.004)	(0.004)	(0.004)
Time FE		X	X	X
Demographics			X	X
Health Controls				X

Notes: Each column lists estimates with standard errors in parentheses (clustered at the zip code level) from separate regressions of neonatal mortality on indicators for delivery in a hospital with No NICU, an Intermediate NICU, and a Community NICU. Regional NICU is the excluded group. The columns successively add controls. Time fixed effects include year dummies, month-of-year dummies, and day-of-week dummies. Demographics include age, age squared, race, ethnicity, and insurance coverage. Health controls include number of prenatal care visits, month in which prenatal care began, parity, sex, multiple birth status, an indicator for having a clinical condition, indicators for small and large for gestational age, and birth weight dummies at 100 gram increments. N = 42,912; * $p<.10$, ** $p<.05$

Table 7: Level of Care by Distance, First Stage Estimates

Dep. Var.:	I(No NICU)				I(Intermediate NICU)				I(Community NICU)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
NoDist	0.116**	0.115**	0.112**	0.111**	-0.044**	-0.044**	-0.042**	-0.043**	-0.034**	-0.031**	-0.033**	-0.034**
	(0.013)	(0.013)	(0.013)	(0.012)	(0.008)	(0.008)	(0.008)	(0.008)	(0.006)	(0.006)	(0.006)	(0.006)
IntDist	-0.087**	-0.087**	-0.085**	-0.084**	0.126**	0.126**	0.124**	0.125**	-0.025**	-0.026**	-0.024**	-0.024**
	(0.015)	(0.015)	(0.015)	(0.014)	(0.013)	(0.014)	(0.013)	(0.013)	(0.009)	(0.009)	(0.009)	(0.009)
ComDist	-0.025**	-0.025**	-0.025**	-0.025**	-0.027**	-0.027**	-0.027**	-0.027**	0.077**	0.073**	0.074**	0.074**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)
F-Stat	33.47	33.22	32.47	32.46	43.53	43.69	44.54	44.56	40.91	37.18	38.48	38.35
Time FE		X	X	X		X	X	X		X	X	X
Demog.			X	X			X	X			X	X
Health				X				X				X

Notes: Each column lists coefficient estimates with standard errors in parentheses (clustered at the zip code level) from separate regressions of a level of care indicator on the distance instruments. See notes to Table 6 for details of control variables. N = 42,912; * p<.10, ** p<.05

Table 8: Neonatal Mortality by Level of Care, 2SLS Estimates

Dependent Variable:	OLS	2SLS
Neonatal Mortality	(1)	(2)
I(No NICU)	0.050**	-0.030
	(0.007)	(0.029)
I(Intermediate NICU)	0.021**	0.009
	(0.005)	(0.015)
I(Community NICU)	0.012**	-0.063*
	(0.004)	(0.035)

Notes: This table lists OLS and 2SLS coefficient estimates with standard errors in parentheses (clustered at the zip code level) of neonatal mortality on indicators for delivery in a hospital with No NICU, an Intermediate NICU, and a Community NICU. Regional NICU is the excluded group. All controls described in the notes to Table 6 are included in both columns. N = 42,912; * p<.10, ** p<.05

Table 9: Robustness - OLS & 2SLS Estimates

	I(No NICU)		I(Intermediate NICU)		I(Community NICU)	
A. Baseline						
OLS	0.050**	(0.007)	0.021**	(0.005)	0.012**	(0.004)
2SLS	-0.030	(0.029)	0.009	(0.015)	-0.063*	(0.035)
B. Zip Code Level Controls						
OLS	0.051**	(0.007)	0.022**	(0.005)	0.012**	(0.004)
2SLS	-0.041	(0.037)	0.008	(0.015)	-0.072*	(0.038)
2SLS (Density Interaction)	-0.003	(0.029)	0.003	(0.013)	-0.027	(0.033)
C. Zip Code Fixed Effects						
OLS	0.055**	(0.007)	0.034**	(0.006)	0.014**	(0.004)
2SLS	-0.186*	(0.108)	0.005	(0.069)	-0.086	(0.056)
2SLS (Density Interaction)	-0.074	(0.098)	0.064	(0.065)	-0.028	(0.045)
D. Pooling No NICU and Intermediate NICU						
OLS			0.032**	(0.004)	0.011**	(0.004)

Notes: Each row lists OLS and 2SLS coefficient estimates with standard errors in parentheses (clustered at the zip code level) of neonatal mortality on indicators for delivery in a hospital with No NICU, an Intermediate NICU, and a Community NICU. Regional NICU is the excluded group. All regressions include all controls described in the notes to Table 6. N = 42,912; * p<.10, ** p<.05

Table 10: Heterogeneity

	Baseline	Hispanic	Non-Black	Medicaid	No College	Suburban	1991 - 1995
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
OLS Estimates							
I(No NICU)	0.050**	0.049**	0.056**	0.042**	0.048**	0.050**	0.059**
	(0.007)	(0.009)	(0.007)	(0.009)	(0.008)	(0.008)	(0.010)
I(Intermediate NICU)	0.021**	0.028**	0.022**	0.018**	0.027**	0.028**	0.017**
	(0.005)	(0.009)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
I(Community NICU)	0.012**	0.010*	0.013**	0.007	0.011**	0.010*	0.013**
	(0.004)	(0.005)	(0.004)	(0.006)	(0.005)	(0.005)	(0.006)
2SLS Estimates							
I(No NICU)	-0.030	-0.023	-0.039	-0.017	-0.018	-0.029	-0.024
	(0.029)	(0.036)	(0.030)	(0.036)	(0.033)	(0.034)	(0.042)
I(Intermediate NICU)	0.009	0.025	0.012	-0.001	0.027	0.022	0.002
	(0.015)	(0.022)	(0.015)	(0.020)	(0.020)	(0.019)	(0.022)
I(Community NICU)	-0.063*	-0.040	-0.080**	-0.047	-0.090**	-0.059	-0.078
	(0.035)	(0.059)	(0.035)	(0.049)	(0.045)	(0.041)	(0.060)
N	42,912	18,974	35,705	21,707	27,651	21,464	19,861
Mean Mortality	0.157	0.168	0.159	0.160	0.162	0.157	0.175

Notes: Each column lists OLS and 2SLS coefficient estimates with standard errors in parentheses (clustered at the zip code level) of neonatal mortality on indicators for delivery in a hospital with No NICU, an Intermediate NICU, and a Community NICU. Regional NICU is the excluded group. All regressions include all controls described in the notes to Table 6. * $p<.10$, ** $p<.05$

Appendix – Not for Publication

This appendix addresses various robustness checks not presented in the paper. I first provide additional tests of the validity of the instruments. I then present specifications estimated with alternative sets of control variables, alternative treatments of the error term, and alternative measures of the mortality dependent variable. I also discuss the first stage estimates behind the specifications presented in the paper with zip code of residence level controls and fixed effects and the pooling of No NICU and Intermediate NICU hospitals. Finally, I show that the main results are not robust to the sample restrictions discussed in the paper. Overall, these additional robustness checks support the finding that OLS estimates overstate the effect of the level of neonatal intensive care available at the delivery hospital on mortality.

A.1 Additional Tests of Instrument Validity

The distance instruments are motivated by the supposition that NICUs are not located according to medical need and are likely adopted to attract low-risk obstetric patients. Table A.1 provides further evidence of this hypothesis by presenting "first stage" estimates of the effect of distance on mothers' hospital choice for infants above the VLBW threshold. I display estimates for low birth weight infants (1,500 to 2,500 grams or 3.3 to 5.5 pounds), those just above the low birth weight cutoff (2,500 to 3,000 grams or 5.5 to 6.6 pounds), and the remaining normal birth weight group (3,000 to 4,500 grams 6.6 to 9.9 pounds).[1] Distances are strong predictors of level of care for these samples, and the coefficient estimates and F-Statistics actually increase in absolute value as birth weight increases. This evidence supports the anecdotes that NICUs attract mothers at all risk levels and the assumption that NICU location is exogenous to the unobserved determinants of VLBW mortality.

I also estimate reduced form regressions of the effect of the distance instruments on neonatal mortality, and examine their sensitivity to the addition of controls. The stability

[1] All samples are subject to the same restrictions described in Section 3 of the paper.

of the first stage estimates discussed in the paper provides evidence in favor of the exclusion restriction. A similar exercise for the reduced form estimates provides a sharper test because it provides evidence on how observable characteristics are correlated with the portion of distance that predicts neonatal mortality. If selection on the unobservables is similar to selection on the observables, and the reduced form estimates are insensitive to controls, the exclusion restriction that $E[\mathbf{D}'_{zt}\varepsilon_{izt}|\mathbf{X}] = \mathbf{0}$ is likely to hold.

Table A.2 presents the results. The first column includes no controls, the second column adds time dummies, and the final two columns add demographic and health characteristics. The estimates are quite stable across specifications. The $NoDist$ and $IntDist$ coefficient estimates change slightly as controls are added, but they are quite small and statistically insignificant across all four columns. The $ComDist$ coefficient estimate is very stable across specifications and in the final column is estimated as a statistically significant -0.004.

Finally, I provide some evidence to support the validity of the instruments when including zip code fixed effects. As discussed in the paper, these specifications assume that changes in unobserved mortality are not correlated with chaneges in differential distance. As in the cross sectional case, this assumption is not formally testable, but Figure A.2 shows that at least trends in mean observable demographic and health characteristics are paralell for zip codes experiencing different changes in distance. It appears that the assumption of parallel trends is reasonable.

A.2 Alternative Control Variables and Clustering

To further test the robustness of my results, Table A.3 presents estimate from six other alternative specifications, with the baseline specification from the paper repeated in Column 1. Columns 2 through 5 test whether the results change of I include various different health related controls. Column 2 adds an indicator for whether the infant was delivered by cesarean section or not. I do not include this control in the main specification because, as a treatment decision, it may be endogenous to the level of neonatal intensive care at the birth hospital. Despite this concern, adding it as a control variable does not appreciably change the OLS or 2SLS estimates. Column 3 and 4 provide evidence that my results are not sensitive to how I control for birth weight. In these two columns, I interact the birth weight indicators with the male dummy and re-specify the birth weight indicators in 50-gram increments instead of 100-gram increments, respectively. Both alternative specifications lead to OLS and 2SLS estimates that are similar to the baseline estimates. Column 5 replaces the dummy indicating whether an infant has any of the defined clinical conditions with a full set of indicators for each of the nine different conditions.[2] Again, the results are quite similar to the baseline estimates.

The last two columns of Table A.3 explore whether the standard error estimates change if the level of clustering is changed. In the paper, standard error estimates are clustered at the zip code level to allow unobserved mortality to be correlated within zip codes. Column 6 allows for more conservative geographic correlation by clustering at the HSA (Hospital Service Area) level. These HSAs are collections of zip codes for which most of their Medicare patients receive care from the same hospital.[3] While these areas are calculated only with Medicare patients, they are likely good proxies for general health care markets. My sample includes 1,144 zip codes which are grouped into 192 HSAs. The standard error estimates remain virtually unchanged when clustering at this larger geographic level. If anything the 2SLS estimate of the community NICU coefficient becomes a bit more precise.[4] Column 7 clusters standard errors by hospital instead of by geography. Allowing unobserved mortality to be correlated within hospitals does slightly inflate the standard errors beyond those allowing unobserved mortality to be correlated within geographic areas.

[2]The nine conditions include hydrops due to isoimmunization, hemolytic disorders, fetal distress, fetus affected by maternal condition, oligohydramnios, other high-risk maternal conditions, placenta hemorrhage, premature rupture of membrane, and prolapsed cord (Phibbs et al. 2007).

[3]Source: `http://gonzo.dartmouth.edu/faq/data.shtm`, last accessed May 17, 2010.

[4]Unreported estimates reveal very similar standard error estimates when clustering at the county level. As a caveat, the asymptotics for clustered standard errors require the number of clusters to approach infinity while the cluster size is fixed. There are only 39 counties in the data, so this specification has a small number of large clusters.

A.3 Alternative Mortality Measures

The main results indicate that OLS estimates overstate differences in neonatal mortality by level of care. This definition of mortality includes all deaths within 28 days of birth or within one year if an infant is continuously hospitalized since birth. It may be the case that results differ for shorter or longer term measures of mortality. In Table A.4 I present OLS and 2SLS estimates of the effect of level of care on 1-day, 28-day, and 1-year mortality, regardless of hospitalization time, repeating the neonatal mortality results in Column 1. In general results are similar to the baseline specification. OLS estimates reveal higher mortality in lower level hospitals. The point estimates increase as the mortality window increases, but increases in the mean mortality rate as the window lengthens imply the relative magnitudes are similar for each outcome. For all three additional mortality outcomes 2SLS estimates are well below the OLS estimates and statistically insignificant. The finding that OLS estimates overstate differences in mortality is robust to these alternative outcome measures.

A.4 Robustness Checks - First Stage Estimates

Section 6 of the paper presents resutls for a variety of alternative specifications. In this section I present the first stage results for the specifications that include zip code level control variables, zip code fixed effects, and the pooling of No and Intermediate NICUs. Results are presented in Table A.5 with the baseline specification from the paper repeated in Panel A. In this table each row lists coefficient estimates from one first stage regression.

When zip code level controls are added in the first three rows of Panel B, the magnitudes of the first stage coefficient estimates change slightly, but they are very similar, highly statistically significant, and the F-Statistics are of similar magnitudes to Panel A. The second portion of Panel B interacts the distance instruments with population density. The coefficient estimates of the three distance measures decrease a bit in magnitude, but are still highly statistically significant. The density interactions are almost all statistically significant with positive diagonal elements and negative off diagonal elements, matching the pattern of signs of the distance main effects. Thus, the effect of distance becomes stronger as population density increases, as would be expected if travel times are longer or travel is more expensive in more densely populated areas. The three added instruments result in similar F-Statistics for the No NICU and Intermediate NICU regressions, but a lower F-Statistic in the Community NICU regression that is still well above 10. Results with zip code of residence fixed effects are presented in panel C. The instruments are still strong predictors of level of care chosen with large, positive, and statistically significant coefficients along the diagonal. The F-Statistics are lower than in the cross sectional specifications, but they are all above 16 without population density interactions and above 11 with the interactions. Panel D indicates that when pooling No NICU and Intermediate NICU hospitals, distance is still a strong predictor of level of care with even larger F-statistics than in the original estimation.

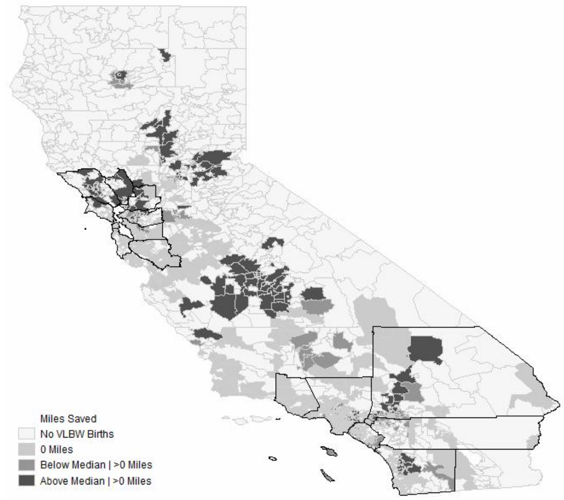
A.5 Sample Selection

In this section I ensure that my estimates are not sensitive to the sample restrictions discussed in Section 3 of the paper. The first column of Table A.6 repeats the estimates from the main specification. Columns 2 through 5 report results including various groups that were excluded from the main analysis sample. Column 2 includes infants in the most rural counties, Column 3 includes infants born in Kaiser hospitals, Column 4 includes infants diagnosed with a congenital anomaly, and Column 5 includes fetal deaths.

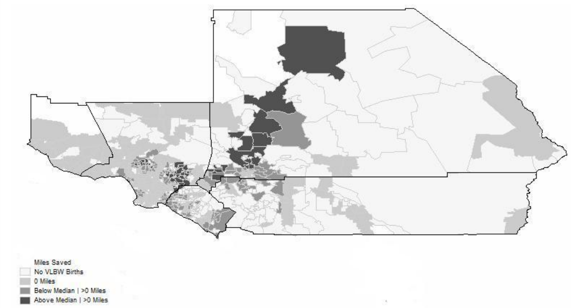
These estimates reveal that the OLS and 2SLS estimates are not appreciably affected by these sample restrictions. If anything, including rural residents results in 2SLS estimates that are closer to zero, although excluding these observations is still probably best, since they are likely to live furthest from all hospitals and may be unobservably different from those living close to all hospitals. Including deliveries in Kaiser hospitals has little effect on the estimation as well. Results of first stage regressions for this sample alone, not shown here, reveal that these added observations do not choose hospitals based on distance; therefore, they do not contribute to the 2SLS estimates, so it is not surprising that the results are not affected by including them.

Including infants with congenital anomalies leads to higher coefficient estimates in the OLS specification, but similar 2SLS estimates to the baseline specification. Finally, including observations of infants who die before delivery approximately doubles the magnitude of both the OLS and 2SLS coefficient estimates. The mean mortality rate for this sample is almost twice that of the main analysis sample, so the relative effects are very similar. This finding indicates that differences in level of care do not differentially impact the probability of death prior to delivery.

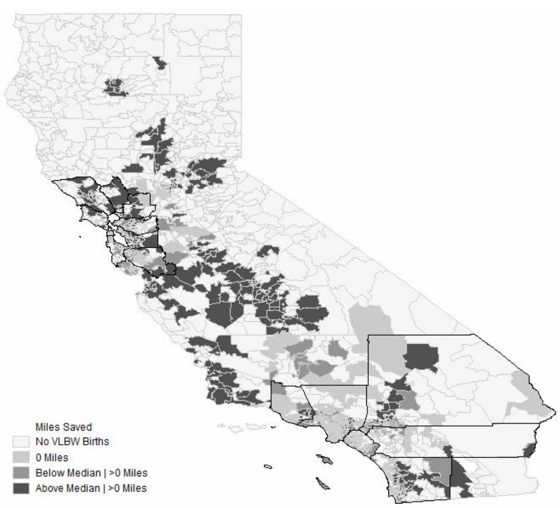
Figure A.1: Miles Saved to Lower Level NICUs, 1991



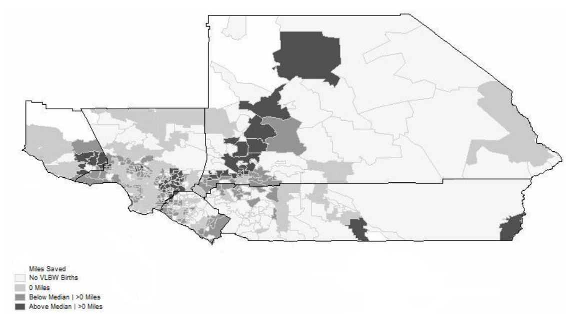
(a) Community Distance: Full State



(b) Community Distance: LA Metro Area



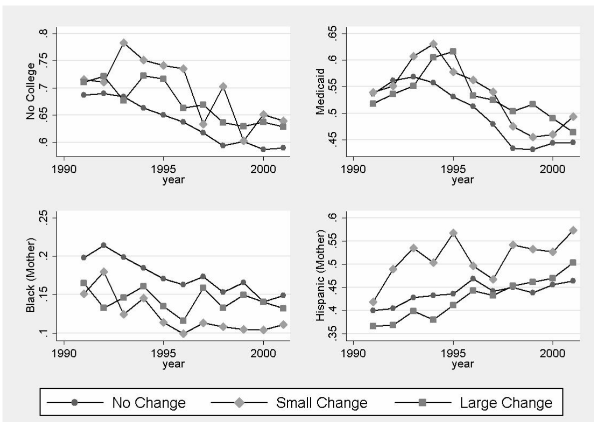
(c) Intermediate Distance: Full State



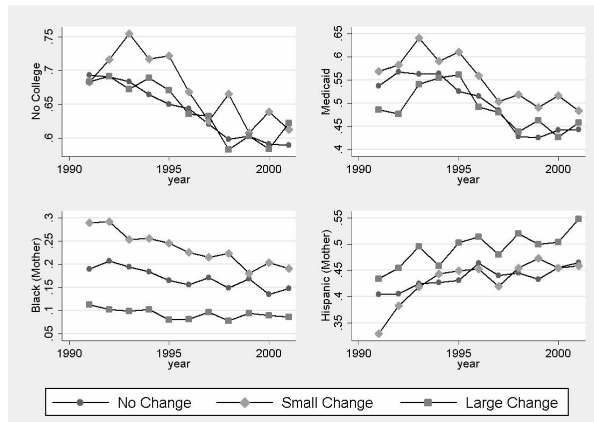
(d) Intermediate Distance: LA Metro Area

Notes: These figures shade zip codes based on the number of miles a mother living at the center of the zip code saves by choosing the nearest Community NICU or higher or the nearest Intermediate NICU or higher over the nearest Regional NICU. Zip codes shaded in white indicate no very low birth weight births in my analysis sample. Remaining zip codes are divided into three groups: those saving zero miles, and those above and below the median conditional on non-zero differential distance. The dark lines in Panel A and Panel C outline counties in the San Fransisco Bay, Los Angeles Metro, and San Diego Metro areas.

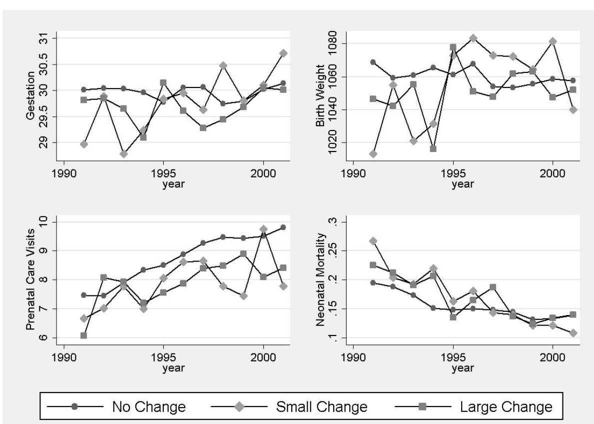
Figure A.2: Demographic and Health Trends by Changes in Distance



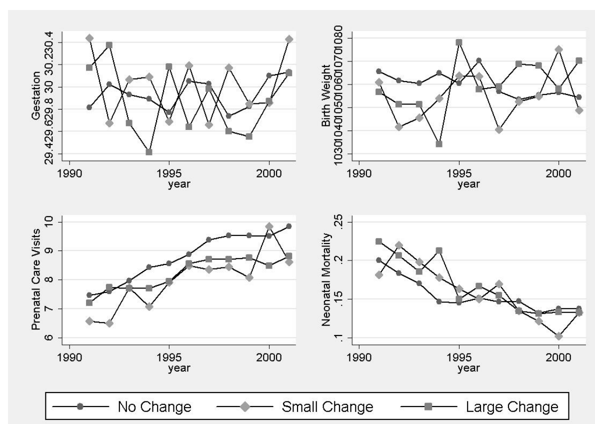
(a) Demographics by Δ IntDist



(b) Demographics by Δ ComDist



(c) Health Characteristics by Δ IntDist



(d) Health Characteristics by Δ ComDist

Notes: These figures plot means of mothers' demographic and infants' health characteristics by changes in differential distance to Intermediate and Community NICUs. Observations are divided into three groups based on whether the zip code of residence becomes no closer, slightly closer (changes below the median), or much closer (changes above the median) to the respective level of care between 1991 and 2001. N=42,912.

Table A.1: Level of Care by Distance for Heavier Infants

	1,500 to 2,500 Grams			2,500 to 3,000 Grams			3,000 to 4,500 Grams		
Dependent Var:	I(No NICU)	I(Inter)	I(Comm)	I(No NICU)	I(Inter)	I(Comm)	I(No NICU)	I(Inter)	I(Comm)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NoDist	0.199**	-0.069**	-0.059**	0.240**	-0.081**	-0.071**	0.251**	-0.083**	-0.077**
	(0.019)	(0.010)	(0.008)	(0.023)	(0.011)	(0.009)	(0.023)	(0.010)	(0.009)
IntDist	-0.173**	0.188**	-0.009	-0.221**	0.209**	0.002	-0.239**	0.209**	0.011
	(0.021)	(0.016)	(0.010)	(0.025)	(0.018)	(0.011)	(0.025)	(0.017)	(0.012)
ComDist	-0.022**	-0.074**	0.086**	-0.015*	-0.087**	0.082**	-0.005	-0.090**	0.080**
	(0.008)	(0.010)	(0.007)	(0.009)	(0.012)	(0.007)	(0.007)	(0.010)	(0.007)
F-Stat	43.27	51.50	57.03	41.99	51.42	59.40	46.84	49.76	64.78
N	237,488	237,488	237,488	751,750	751,750	751,750	3,705,006	3,705,006	3,705,006

Notes: Each column lists coefficient estimates with standard errors in parentheses (clustered at the zip code level) from separate regressions of delivery hospital level of care indicators on No NICU distance, Intermediate distance, and Community distance. Each panel presents estimates for a different sample, stratified by birth weight. All regressions include all controls described in the notes to Table 6 of the paper. * p<.10, ** p<.05

Table A.2: Reduced Form Estimates

	Dependent Variable: Neonatal Mortality			
	(1)	(2)	(3)	(4)
NoDist	0.003	0.002	-0.002	-0.002
	(0.004)	(0.004)	(0.004)	(0.003)
IntDist	0.002	0.002	0.004	0.005
	(0.005)	(0.005)	(0.005)	(0.003)
ComDist	-0.005**	-0.003	-0.003	-0.004**
	(0.003)	(0.003)	(0.003)	(0.002)
Time FE		X	X	X
Demographics			X	X
Health Controls				X

Notes: Each column lists OLS coefficient estimates with standard errors in parentheses (clustered at the zip code level) from separate regressions of neonatal mortality on No NICU distance, Intermediate distance, and Community distance. Columns successively add controls. See notes to Table 6 of the paper for details of control variables. N = 42,912; * $p<.10$, ** $p<.05$

Table A.3: Alternative Control Variables and Clustering

	Baseline	C-Section Control	BW Dummies × Male	BW Dummies in 50 Grams	Specific Clin. Cond. Controls	Cluster at HSA	Cluster at Hospital
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
OLS Estimates							
I(No NICU)	0.050**	0.049**	0.049**	0.050**	0.049**	0.050**	0.050**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
I(Intermediate NICU)	0.021**	0.020**	0.021**	0.020**	0.020**	0.021**	0.021**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.009)
I(Community NICU)	0.012**	0.013**	0.011**	0.012**	0.012**	0.012**	0.012*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.007)
2SLS Estimates							
I(No NICU)	-0.030	-0.028	-0.030	-0.025	-0.030	-0.030	-0.030
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.030)	(0.034)
I(Intermediate NICU)	0.009	0.009	0.010	0.009	0.010	0.009	0.009
	(0.015)	(0.015)	(0.014)	(0.014)	(0.015)	(0.015)	(0.023)
I(Community NICU)	-0.063*	-0.058*	-0.063*	-0.057*	-0.060*	-0.063**	-0.063
	(0.035)	(0.035)	(0.034)	(0.034)	(0.035)	(0.029)	(0.043)

Notes: Each column lists OLS and 2SLS coefficient estimates with standard errors in parentheses (clustered at the zip code level) of neonatal mortality on indicators for delivery in a hospital with No NICU, an Intermediate NICU, and a Community NICU. Regional NICU is the excluded group. All regressions include all controls described in the notes to Table 6 of the paper plus additional controls noted in column headings. N = 42,912; * p<.10, ** p<.05

Table A.4: Alternative Mortality Measures

	Neonatal Mortality	Mortality within: 1 Day	28 Days	1 Year
	(1)	(2)	(3)	(4)
OLS Estimates				
I(No NICU)	0.050**	0.026**	0.048**	0.052**
	(0.007)	(0.005)	(0.006)	(0.007)
I(Intermediate NICU)	0.021**	0.017**	0.018**	0.024**
	(0.005)	(0.004)	(0.005)	(0.006)
I(Community NICU)	0.012**	0.003	0.011**	0.012**
	(0.004)	(0.003)	(0.004)	(0.004)
2SLS Estimates				
I(No NICU)	-0.030	-0.054**	-0.030	-0.014
	(0.029)	(0.020)	(0.026)	(0.027)
I(Intermediate NICU)	0.009	0.011	-0.003	0.021
	(0.015)	(0.011)	(0.014)	(0.014)
I(Community NICU)	-0.063*	-0.099**	-0.097**	-0.037
	(0.035)	(0.027)	(0.033)	(0.035)
Mean Mortality	0.157	0.063	0.140	0.170

Notes: Each column lists OLS and 2SLS coefficient estimates with standard errors in parentheses (clustered at the zip code level) of neonatal mortality on indicators for delivery in a hospital with No NICU, an Intermediate NICU, and a Community NICU. Regional NICU is the excluded group. All regressions include all controls described in the notes to Table 6 of the paper. * p<.10, ** p<.05

Table A.5: Robustness - First Stage Estimates

	NoDist		IntDist		ComDist		NoDistXDensity		IntDistXDensity		ComDistXDensity		F-Stat
A. Baseline													
I(No NICU)	0.111**	(0.012)	-0.084**	(0.014)	-0.025**	(0.007)							32.46
I(Intermediate)	-0.043**	(0.008)	0.125**	(0.013)	-0.027**	(0.009)							44.56
I(Community)	-0.034**	(0.006)	-0.024**	(0.009)	0.074**	(0.008)							38.35
B. Zip Code Level Controls													
I(No NICU)	0.105**	(0.012)	-0.079**	(0.014)	-0.025**	(0.007)							29.83
I(Intermediate)	-0.038**	(0.007)	0.120**	(0.013)	-0.026**	(0.009)							44.01
I(Community)	-0.045**	(0.008)	-0.015	(0.010)	0.073**	(0.008)							38.30
I(No NICU)	0.097**	(0.012)	-0.066**	(0.014)	-0.025**	(0.008)	0.020**	(0.005)	-0.022**	(0.005)	0.002	(0.002)	26.60
I(Intermediate)	-0.023**	(0.007)	0.080**	(0.013)	-0.010	(0.011)	-0.011**	(0.004)	0.034**	(0.009)	-0.020**	(0.008)	46.54
I(Community)	-0.046**	(0.007)	0.009	(0.010)	0.053**	(0.008)	-0.007	(0.006)	-0.012*	(0.007)	0.018**	(0.005)	17.92
C. Zip Code Level Fixed Effects													
I(No NICU)	0.085**	(0.016)	-0.111**	(0.017)	-0.009	(0.006)							16.06
I(Intermediate)	-0.000	(0.025)	0.124**	(0.017)	-0.045**	(0.005)							46.08
I(Community)	0.021	(0.036)	0.074**	(0.028)	0.072**	(0.011)							21.50
I(No NICU)	0.075**	(0.019)	-0.108**	(0.018)	-0.005	(0.006)	0.012**	(0.005)	-0.006	(0.006)	-0.005	(0.003)	11.92
I(Intermediate)	-0.001	(0.025)	0.107**	(0.020)	-0.041**	(0.005)	-0.008	(0.005)	0.021**	(0.007)	-0.005	(0.005)	28.94
I(Community)	-0.004	(0.037)	0.048*	(0.028)	0.060**	(0.011)	-0.002	(0.011)	0.029**	(0.012)	0.012*	(0.007)	13.61
D. Pooling No and Intermediate NICUs													
No/Interm	0.093**	(0.007)			-0.037**	(0.008)							114.74
Community	-0.048**	(0.005)			0.065**	(0.006)							60.04

Notes: Each row lists coefficient estimates with standard errors in parentheses (clustered at the zip code level) from separate regressions of delivery hospital level of care indicators on No NICU distance, Intermediate distance, and Community distance. All regressions include all controls described in the notes to Table 6 of the paper. N = 42,912; * p<.10, ** p<.05

Table A.6: Robustness to Sample Restrictions

	Baseline	Include Rural	Include Kaiser	Include Cong. Anom.	Include Fetal Death
	(1)	(2)	(3)	(4)	(5)
OLS Estimates					
I(No NICU)	0.050**	0.054**	0.052**	0.058**	0.110**
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)
I(Intermediate NICU)	0.021**	0.023**	0.026**	0.029**	0.056**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
I(Community NICU)	0.012**	0.013**	0.013**	0.011**	0.023**
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
2SLS Estimates					
I(No NICU)	-0.030	-0.006	-0.017	-0.000	-0.066**
	(0.029)	(0.028)	(0.028)	(0.029)	(0.028)
I(Intermediate NICU)	0.009	0.005	0.013	0.018	0.014
	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)
I(Community NICU)	-0.063*	-0.020	-0.055*	-0.050	-0.093**
	(0.035)	(0.033)	(0.033)	(0.036)	(0.036)
N	42,912	44,937	49,018	47,121	51,123
Mean Mortality	0.157	0.157	0.158	0.167	0.291

Notes: Each column lists OLS and 2SLS coefficient estimates with standard errors in parentheses (clustered at the zip code level) of neonatal mortality on indicators for delivery in a hospital with No NICU, an Intermediate NICU, and a Community NICU. Regional NICU is the excluded group. All regressions include all controls described in the notes to Table 6 of the paper. * p<.10, ** p<.05