APPENDIX D: Identifying Hospital Mergers

The American Hospital Association Annual Survey contains data on respondent hospitals in the United States. While the AHA data are an invaluable source of information regarding hospital characteristics and geography, they provide an incomplete picture of hospital ownership transitions for multiple reasons. As a result, we have implemented several corrections in order to create more comprehensive and accurate roster of hospital mergers.

First, the AHA reports data for a reference period preceding the year of the survey. As a result, system information in the AHA typically refers to the year following the reference year. In general, we deal with this issue by utilizing the lagged system information.

Second, the AHA sometimes deals with mergers and acquisitions in a way that complicates our analysis. In a case where one hospital merges with another, the AHA contains a single observation for the merged entity. However, when a system *acquires* a hospital, it sometimes retains its unique AHA ID and experiences a change in its system ID. In order to obtain a complete picture of hospital geography and ownership, we generate imputed observations for those hospitals, which are deleted from the data as a result of a merger, while noting the change in ownership structure through the system information.

We also incorporate several additional changes, which are motivated by validations with external data sources. We used data from the following databases to track mergers and acquisitions of hospitals: Irving-Levin Associates, Factset, and SDC Platinum. Each database contains detailed information (e.g. parties involved, announcement and closing dates) on both completed and failed M&A deals. To incorporate this merger information into the AHA survey, we aggregated the 2006 to 2011 AHA surveys to create a panel dataset of hospital IDs where the time unit is year. We then created a new health system ID for each hospital (called *sysid2*). If a hospital's health system ID (*sysid* in the AHA survey) was non-missing, we assigned this health system ID to *sysid2*. For those hospitals where the value was missing, we "filled-down" *sysid2* with the health system ID of the first non-missing year before it. That is, if hospital "A" had health system ID "B" in 2007, but the *sysid* was blank in 2008 and 2009, we assigned health system "B" to "A" in 2008 and 2009.[13] If a hospital had a missing *sysid* for all years, then we assigned the hospital's unique hospital ID number to the *sysid2* to denote that the hospital was an independent hospital in all years.

Next, we reviewed each merger description in the three M&A databases and determined which AHA hospital IDs were parties to a merger, which health system IDs corresponded to the parties involved, and the date the merger closed. We then recorded the system ID of the acquiring party in a new variable called *ilsysid* for the target hospital in the year the merger closed.[14] After completing this for every hospital merger in the three M&A databases, we then "filled-down" the blank values of the *ilsysid* variable in a similar manner to *sysid2* (i.e. the years where a hospital did not experience a merger). If the hospital was not involved in any mergers (according to our

[13] "Fill-down" in this context assumes the panel data are sorted by AHA hospital ID number and in ascending order by year.

[14] If the closing date was not populated in one of the M&A databases and we could not find a news article or report that documented the closing of the deal, we used the announcement date as an estimate of when the deal was completed.

three databases), then we assigned the hospital's unique ID number to the *ilsysid* variable to denote it was an independent hospital.

We then flagged all instances where *sysid2* did not equal *ilsysid* and reviewed each instance on a case-by-case basis to determine why there were discrepancies between the two health system IDs. We used resources such as Becker's Hospital Review and local newspapers to determine if *sysid2* or *ilsysid* (or neither) were the correct health system ID. We then created a consolidated health system ID variable (called *msysid*) to account for this new information; *msysid* is the variable we use to identify mergers. If the *msysid* of hospital h switched from s_1 to s_2 between year $t-1$ and year t, then we say hospital h experienced a merger in year t.

We measure the distance between merging parties using straight-line distance between hospitals. In classifying mergers, we define a target (acquired) hospital as experiencing a merger within X miles if there is at least one hospital in the acquiring system located within X miles. Symmetrically, only the members of the acquiring system that are within X miles of the target are considered "X mile" mergers.

We have created a database of hospital mergers that are available at www.healthcarepricingproject.org.