Supplemental Appendix to: "Consolidation in the Dialysis Industry, Patient Choice, and Local Market Competition"

Calculating the Herfindahl-Hirschman Index (HHI):

The equation used to calculate HHI is as follows:

$$HHI_{hospital\ service\ area} = \sum_{i=1}^{n} s_i^2$$

Where S_i represents the proportion of patients living in an HSA receiving dialysis at the i^{th} firm in the HSA.

We calculate this measure of HHI for each hospital service area (HSA) from the following three steps:

1. Calculate a "first-stage" competition measure for each HSA (using the equation above) based on the market shares of the firms where patients living in each HSA choose to dialyze. In this stage all patients residing in a given HSA define the "market" for each firm-HSA pair. Firms do not have to be located in the same HSA where patients reside, and a given dialysis facility can be included in the calculation of HHI for multiple HSAs if patients from multiple HSAs dialyze at that facility.

2. Calculate a "second-stage" dialysis-facility-level measure of competition using a weighted average of the "first-stage" HSA-level HHIs for patients who actually dialyze at each facility. This measure is calculated for each separate facility, regardless of which organization owns a facility, and assumes facilities can discriminate between patients living in different HSAs (or are aware of any differences in choices available to patients living in different HSAs) when competing against rival providers. This step is necessary in order to convert the HHI calculated from step 1 above to a measure of competition that dialysis facilities face.

3. Calculate a final ("third-stage") HSA-level measure of competition from a weighted average of the facility-level-HHIs at facilities where patients residing in each HSA receive dialysis.

In summary, this index represents a weighted average of competition indices for facilities that treat patients in a given HSA, where facility competition indexes are, in turn, weighted by choices available to patients they treat.

Comorbidity Index:

A modified Charlson score was obtained using the following algorithm:

1) 1 point was assigned to the following comorbidities: peripheral vascular disease, dementia, chronic lung disease, rheumatological disease, peptic ulcer disease, diabetes.

2) 2 points were assigned to the following conditions: myocardial infarction, congestive heart failure, cerebral vascular disease, diabetes, liver disease, cancer.
3) 6 points were assigned to patients with HIV

Cerebrovascular disease included diagnoses of central nervous system bleed, central nervous system vascular disease, stroke, or transient ischemic attack. Peptic ulcer disease included patients with diagnosed peptic ulcer disease along with diagnosed gastrointestinal bleed. Myocardial infarction included patients diagnosed with an acute myocardial infarction, a history of a coronary bypass or percutaneous coronary intervention, and a history of unstable angina. Heart failure included patients with diagnosed heart failure along with valvular disease. Chronic lung disease included patients with diagnosed lung disease or pulmonary hypertension. We assumed a diagnosis of liver disease was moderate or severe.

In an effort to retain as much information as possible, we combined some elements of the modified Charlson comorbidity index with the original Charlson comorbidity index. Specifically, our calculation deviated from the modified Charlson morbidity index described by Hemmelgarn et al. in the following ways[1]:

1) Since we did not have reliable information on the type of cancer, we assigned 2 points to a diagnosis of cancer (which is what the original Charlson comorbidity index assigns for neoplasia), and did not include separate point scores for lymphoma or metastatic disease.
2) We included 6 points for HIV, which is what is assigned in the traditional Charlson comorbidity index (the modified index did not include HIV due to insufficient data).
3) We assigned 2 points for hemiplegia, which is what is used in the original Charlson comorbidity index (the modified index did not include hemiplegia due to insufficient data).

Sensitivity Analysis:

Sensitivity to use of zip code centroids and assumption that the distance is zero for patients living in the same zip code as their dialysis facility.

In our primary analysis of choices among competing dialysis facilities, we made two assumptions when calculating the distance between patient residences and neighboring dialysis facilities. First, we assumed that patients lived at the centroid of their zip code. Additionally, we assumed that the distance traveled was zero in instances where patients lived in the same zip code as their dialysis facility. We made these assumptions in order to avoid bias, despite the fact that they may misrepresent actual distances in certain cases.

There are two instances, in particular, where these assumptions are likely to be less representative of actual distances between patients and nearby facilities. The assumption that patients live at the centroid of their zip code may have led us to assign distances between patients and facilities that are substantially different from actual distances in more remote areas where zip codes can cover a long distance. The assumption that the distance is zero when patients share the same zip code as a dialysis facility could mischaracterize what are actually substantial travel distances in more densely populated areas where travel may be difficult.

We tested the sensitivity of our analysis of the change in number of choices available to patients within their choice radius between 2001 and 2011 to these assumptions in the following steps:

1) We obtained data on land areas and water areas for each census-based zip code tabulation area (based on the 2010 census) from a published dataset.[2] We calculated the total area for each zip code tabulation area by adding the reported land and water areas.
2) We estimated distances from the centroid of a zip code tabulation area to its periphery by taking the average distance obtained under two scenarios:
    a. In one scenario, we assumed that zip code areas are squares. We calculated the distance from the center to the periphery of each zip code tabulation area using the following equation:

$$distance = \sqrt{total\ area}/_2$$

    b. In a second scenario, we assumed that zip code tabulation areas are circles. We calculated the distance from each zip code centroid to its boundary using the following equation:

$$distance = \sqrt{total\ area/_\pi}$$

3) We merged zip code tabulation areas to a complete list of zip codes and then obtained the average distances from zip code centroids to the periphery of zip codes within the following four population density categories: Metropolitan, Micropolitan, Small Town, Rural. This yielded the following four average distances:

Population density category	Mean distance to zip code periphery (miles)
Metropolitan	2.55
Micropolitan	4.61
Small Town	5.32
Rural	5.34

4) For each patient in our cohorts in 2001 and 2011, we obtained an "adjusted" distance to neighboring facilities by assuming that patient residences were uniformly distributed within their zip codes along a line from the centroid to the periphery in each direction from the centroid. From this assumption, we used random sampling to assign a distance adjustment to each patient using the following method:
    a. For each patient (and year), we sampled from a uniform distribution with a minimum of -1.0 and maximum +1.0. We then multiplied this randomly-generated number by the mean distance from zip code centroid to periphery

(listed in the above table) among patients living in similar population densities to obtain a distance adjustment for each patient.

b. For each patient (and year) we added the distance adjustment to the distance calculated in our primary analysis to obtain an adjusted distance to each nearby dialysis facility used in the sensitivity analysis.
c. We used a slightly different method for calculating adjusted distances among patients residing in the same zip code as their dialysis facility (where the distance was zero in our primary analysis). In these instances, the distance could not become less than zero. Additionally, both dialysis facilities and patients could move from the centroid. Consequently, for these patients, we:
    i. Multiplied the mean distance from zip code centroid to periphery (illustrated in the above table) by two.
    ii. Drew from a random uniform distribution ranging from zero to one.
    iii. Calculated the adjusted distance by taking the product of steps 'i' and 'ii' above.

5) For each patient, we used the adjusted distances to calculate the number of facilities within their choice radius in 2001 and in 2011, and the change in facilities within their choice radius.
6) We performed this calculation in 10 different simulations, drawing randomly from uniform distributions for each patient in each simulation, and took the average change in facilities available from the 10 simulations.

It is important to note that we did not, in this sensitivity analysis, use adjusted distances to calculate a new choice radius for each HSA. The choice radii were derived from our primary data. Consequently, this sensitivity analysis focuses on the sensitivity of our estimate of the *change* in the number of facilities available for patients to choose from within their choice radius.

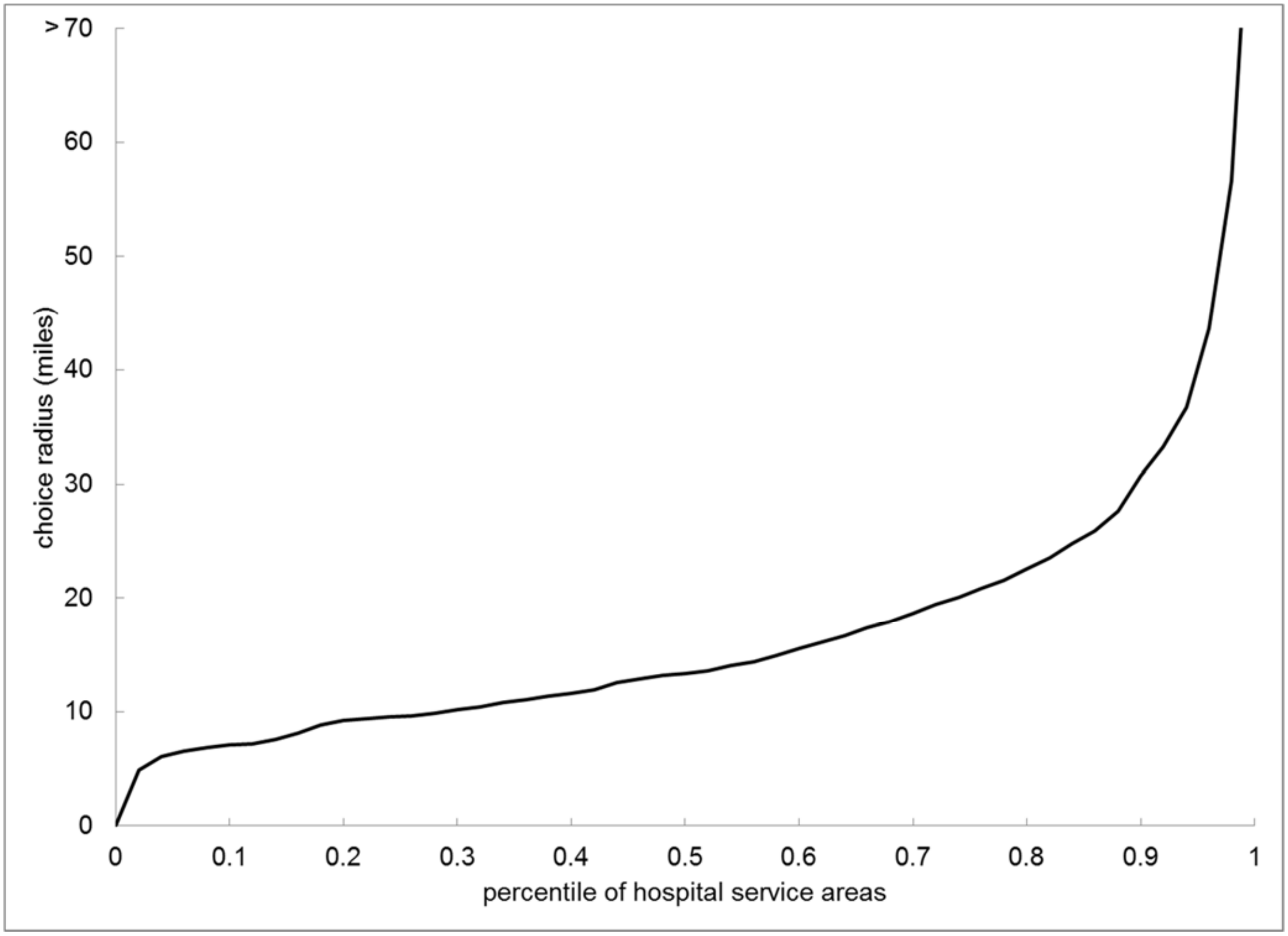
We found that our estimate of the number of competing facilities available to choose from was not sensitive to our use of zip code centroids and our assumption that travel distance was zero among patients living in the same zip code as their dialysis facility. In particular, the sensitivity analyses found that patients, on average, had 7.0 competing facilities in their choice radius in 2001, and that this increased to 7.6 competing facilities in 2011. This was virtually identical to the findings from our primary analysis (an increase from 6.9 in 2001 to 7.6 in 2011).

Table S1: Large Mergers in the Dialysis Industry between 2001 and 2011:

Year of Merger/Acquisition	**Acquiring Organization**	**Acquired Organization**	**Number of Facilities Acquired[1]**
2004	Renal Care Group	National Nephrology Associates	87
2004	DaVita	Gambro Healthcare	560
2005	Fresenius Medical Care	Renal Care Group	425

[1] Based on the initial press release

Figure S1: Distribution of Choice Radii in 2001



Note: Choice radius refers to the distance (in miles) within which 90 percent of patients in a hospital service area travel to receive hemodialysis. Data are plotted for every 2 percentile points. Hospital Service Areas are weighted by the number of patients living in them.

REFERENCES:

1. Hemmelgarn BR, Manns BJ, Quan H, Ghali WA. Adapting the Charlson Comorbidity Index for use in patients with ESRD. *American Journal of Kidney Diseases.* 2003;42:125-132.
2. ProximityOne. Census 2010 ZIP Code (ZCTA) Characteristics In: Dataset CZCDP, ed2016.