Supplemental Digital Content 1. Technical Appendix

In this appendix, we provide further details on how we constructed our concentration measure as well as details on our regression analysis.

Concentration Measure

We constructed Hirschman-Herfindahl Indices (HHIs) for anesthesia groups. By convention, HHIs have a maximum value of 10,000, reached in monopoly markets. As the amount of competition increases, the HHI falls and approaches 0 as the number of practices increases and the size of each individual practice falls.

Computing HHIs requires defining product markets. Here we take product markets to include all anesthetics, which are identified using the relevant Current Procedural Terminology CPT codes. HHIs also require defining geographic markets. We derive geographic markets for each practice empirically, based on observed patient flows in the claims data. This approach seems superior to approaches that would identify HHIs under the assumption that practices serve areas with boundaries defined by larger areas such as counties—as discussed in the main paper, to the degree that anesthesia groups compete across county lines, using a fixed geographic boundary such as a county will tend to overestimate the degree of concentration. For example, suppose two groups are located in neighboring counties (counties A and B) and do compete against each other for contracts from the hospitals in those counties. By happenstance, group A provides all the anesthetics in county A and group B provides all the anesthetics in county B. Since the two groups compete for the contracts located in both counties, the true market is counties A and B combined, and each group has a 50% market share. However, a simple analysis at the county level would treat the two counties as separate markets and assign each group a 100% share in their respective county.

Our analytic approach adapts the approach of Kessler and McClellan to the case of hospitals.[24] We derive HHIs for in two steps. We begin by constructing a ZIP code HHI for each ZIP code, by specialty, by year. Denote by $service_{i,j}$ the number of claims provided by physicians in practice *i* to patients who reside in ZIP code *j* . Denoting the total number of claims provided to patients in ZIP *j* as $service_j$, the market share of practice *i* for ZIP *j* is $share_{i,j} = service_{i,j} / service_j$. The ZIP code HHI is then the sum of squared market shares:

$$ZIPHHI_j = \sum_{\substack{practices\ i \\ serving\ ZIP\ j}} share_{i,j}^2$$

This construction allows flexibility in the market size, basing the HHI on the set of physicians actually observed to provide services to patients in the given ZIP code. We exclude from this calculation claims where the physician is more than 100 miles from the patient ZIP, to reduce the potential for bias from cases where a patient, perhaps while traveling, sees a distant physician who does not play a substantial role in competition for patients residing in the ZIP code. (Distances were determined based on the centroid of the patient and provider ZIP codes, using the Haversine formula. Between 90 and 95% of claims meet the 100 mile criteria in any given year.)

In the second step, for each practice we identified the set of patient ZIP codes with non-zero service claims (i.e. the set of *j* for which $service_{i,j} > 0$), excluding cases where the patient ZIP is more than 100 miles from the physician ZIP. Following the FTC/DOJ guidance, we then took the smallest set of these ZIP codes that accounted for 75% of allowed charges as the practice market area for analysis. We averaged the ZIPHHI values for the ZIP codes in the market area, weighting by the number of services practice *i* provides in each of the patient ZIPs in its market area, to create a practice level HHI:

$$PRACHHI_i = \sum_{\substack{ZIPs\ j\ in \\ market\ area \\ of\ practice\ i}} w_{i,j} ZIPHHI_j$$

where $w_{i,j}$ is a weight with sum 1 derived from the $service_{i,j}$ values (i.e. $service_{i,j}$ / $service_i$ where $service_i$ is the sum of all claims provided by practice *i*).

This approach diverges somewhat from approaches that would simply define the market area of the practice as the set of ZIP codes served and then compute the HHI from the market shares of all practices serving the area. Our approach allows us to increase the weight put on areas from which the practice draws most of its patients. Many practices draw patients from a large number of ZIP codes in total, but have a much smaller set of areas from which the bulk of their patients come. Weighting by the concentration of patients should make the HHIs more accurate in this sense.

For analysis, we created county level measures of the average PRACHHI of physicians located in the county. Denoting areas by k, we take the average of PRACHHI values over the practices *i* with provider locations in county k, weighting by the services provided by the practice attributable to area k.

$$GEOHHI_k = \sum_{\substack{practices\ i \\ with\ provider \\ ZIPs\ in\ area\ k}} b_{i,k} PRACHHI_j$$

where *b* is a weight that sums to one, capturing the distribution across practices of claims attributable to county *k* (i.e. $b_{i,k} = service_{i,k} / service_k$). The principle of weighting here is to upweight practices that have a prominent presence in the area, and downweight practices that do not.

To examine robustness to alternate specifications, we computed HHIs in a number of different ways. We examined the effects of 1) using the number of claims as the service unit; 2) using the number of work RVUs as the service unit; 3) using all ZIPs, rather than the subset accounting for 75% of allowed charges, as the service area; and 4) relaxing the restriction that the physician and patient must be within 100 miles for the claim to be included. All of these produced very similar results, with county-level correlations of 0.97 or higher within each specialty. Finally, we computed HHIs using the SK&A data linked to the Medicare claims, using the largest reported practice entity for each physician (SK&A data allow physicians to designate a group practice, a hospital owner, a system owner, or any combination of those 3. From that, we took the practice size with the most other physicians). The correlations between the specialty median HHIs based on TIN and SK&A group code are very high -- 0.98 or higher across specialties.

Regression Model

For our study, we used a difference-in-differences approach to identify the effect of concentration on payments for anesthesia CPTs. We implemented our difference-in-differences approach using the following regression:

$$\boldsymbol{ln(payment)_{ijt} = f_i + g_j + \delta_t + trend_{it} + \Gamma X_{it} + d_{it}^{25-50} + d_{it}^{50-75} + d_{it}^{75-100} + \varepsilon_{ijt}}$$

In the equation above, $\boldsymbol{payment_{ijt}}$ is the mean payment for CPT *j* in county *i* at year *t*. $\boldsymbol{f_i}$represents a fixed effect for county *i* , $\boldsymbol{\delta_t}$ represents a year effect for year *t,* and $\boldsymbol{g_j}$ represents a fixed effect for CPT *j.* $\boldsymbol{trend_{it}}$ controls for linear trends at the county level and is the product of the year and an indicator variable for the given county. $\boldsymbol{X_{it}}$ is a vector of county characteristics, including total population, percentage of the population that is white, percentage of the population that is over age 65, percentage of the population that is male, and median county income. The set of dummy variables $\boldsymbol{d_{it}}$ indicate whether the county's HHI lies in the 25^{th}-49^{th} percentile ($\boldsymbol{d_{it}^{25-49}}$), 50-$74^{th}$ percentile ($\boldsymbol{d_{it}^{50-74}}$), and 75^{th}-100^{th} percentile ($\boldsymbol{d_{it}^{75-100}}$). Since the omitted group is the 0-24^{th} percentile, the value of these coefficients represents the increase in payments associated with the given quartile relative to this group. $\boldsymbol{\varepsilon_{it}}$ represents the error term.

In the regression above, note that we use the natural log of payments as our independent variable. We do so because of the wide range in payments across anesthesia CPTs (see table 1). In this context, the coefficients on the dummy variables $\boldsymbol{d_{it}}$ can be translated into the percentage change in payment (relative to the 0-25th percentile) associated with the given quartile using methods described elsewhere.[30] Note, however, that generally the coefficient itself approximates the change in payment associated with the given quartile. For example, if the coefficient associated with $\boldsymbol{d_{it}^{75-100}}$ is 0.12, this would generally mean that payments for counties in the 75th to 100th percentile of HHI are 12 percent higher than payments in counties in the 0 to 25th percentile.

Our dataset consists of 10,305 observations (229 counties times 9 years times 5 procedures per county-year). A simple ordinary least squares (OLS) regression will tend to underestimate our standard errors (and therefore overestimate the statistical significance of our regression coefficients) because the 10,305 observations are not likely to represent truly independent observations; observations within a given county are likely to be correlated. Calculating clustered standard errors is an appropriate approach to deal with this issue.[a] In essence, clustering adjusts the standard errors based on the observed level of correlation within a given unit (cluster) defined by the investigator. Since we are primarily concerned with correlation within a given county (across CPT codes and over time), we cluster our standard errors at the county level.[b] Many statistical packages can easily calculate clustered standard errors.

Heteroscedasticity is also potential issue in our regression, since our dependent variable is an average of random variables. Moreover, our independent variable of interest—our

[a] See Wooldridge JM: Cluster-Sample Methods in Applied Econometrics. The American Economic Review 2003; 93: 133-138

[b] See Bertrand M, Duflo E, Mullainathan S: How Much Should We Trust Differences-in-Differences Estimates? The Quarterly Journal of Economics 2004; 119: 249-275

measure of concentration—is also, as described above, an average of random variables. We address the potential for heteroscedasticity in two ways. First, we perform a weighted least squares regression, where the weights are the underlying number of claims used to calculate the average payment in a given county. Second, we calculate clustered standard errors, which—akin to Huber-White errors—are robust to heteroscedasticity.[c]

[c] See Petersen M: Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. Review of Financial Studies 2009; 22: 435-480