

## **A. Collection of Information Employing Statistical Methods**

We will randomly sample Medicare claims to select the providers that will be surveyed. We will stratify our sample by rural and non-rural localities. We will further stratify to assure that less populous states (the states at or below the 25<sup>th</sup> percentile population) are fully represented. Finally, we will stratify to assure an appropriate balance of areas with high, low and average consumer prices (our proxy indicator for this purpose will be the Medicare Geographical Practice Cost Index (GPCI) for the locality). We will stratify at the 1<sup>st</sup>- 25<sup>th</sup> percentile, the 26<sup>th</sup> – 75<sup>th</sup> percentile and the 75<sup>th</sup>- 100<sup>th</sup> percentile of the range of GPCIs. Within each stratum, we will sample based on the last digits of the health insurance claim number (which generally correspond to the Social Security number of the beneficiaries) to achieve the sample size desired (e.g., for a 2% sample we would select two numbers between 00 and 99 and then use claim numbers ending with those two numbers).

Our sample size will be selected with the intent of achieving mean results with 2% error rates and 95% confidence. We will use the mean and standard deviation of submitted charges to as our initial estimate for the associated cost variance to choose our sample size. As we develop trends for items and services, we will revise our estimated variances accordingly.