Data Supplement

In this study, we use a difference-in-difference approach, comparing ED use before and after implementation of the alternative quality contract (AQC). Enrollees in the intervention group were those with primary care providers (PCPs) participating in the AQC in 2009 and the control group were enrollees with PCPs not enlisting in the AQC. This basic setup allows us to relate changes in ED use (our dependent variable) to the implementation of the AQC in a linear ordinary-least-squares model. We controlled for age, sex, enrollee risk scores, and time. Co-morbidities were accounted for in the risk score, which was calculated for each enrollee by Blue Cross Blue Shield (BCBS) of Massachusetts. The score is similar to the diagnostic-cost group (DxCG) scoring system (Verisk Health, Eden Prairie, MN). It is derived by statistical analyses using national claims data, current-year diagnoses, and demographic information. The DxCG is similar to Medicare's Hierarchical Condition Category system, but more detailed and designed for younger persons. We controlled for time non-parametrically in our base model, using dummy variables each quarter. Modifications to our statistical model are summarized below.

Several sensitivity analyses were performed. First, the analysis was restricted to 548,677 individuals with continuous enrollment for 48 months. Second, the linear risk score was redefined into boundaries by its 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th}, and 95^{th} percentiles, and models were rerun. Third, secular trends were controlled for using a linear time variable rather than quarter dummies. Last, all models were repeated without propensity score weighting for comparison.

The findings from our sensitivity analyses support our main findings. Restricting our analysis to those continuously enrolled for 48 months did not alter our estimates. This demonstrated that enrollee entry or exit between plans or provider groups did not affect the AQC estimates during the time of our analysis. Relative utilization by risk category remained the same

after evaluation by quintiles and tertiles. Using the alternative time variables instead of quarter dummy variables did not change the findings. Omitting the propensity score weights used to account for the lack of randomization and balance individual characteristics between groups did not affect results. Comparison of the enrollees in each group showed that individuals were similar in age, sex, and risk scores between the intervention and control groups.

Regression Model to Identify the Impact of the AQC on ED Use

Y = b1*__pre-post__ + b2*__AQC-control__ + b3*__d-in-d__ + b4*__age__ + b5*__age__*__sex__ + b6*__risk_score__ + b6*__qtr__ + b7*__qtr*AQC__ + b9*__propensity_score__

pre-post = pre-post intervention indicator; **AQC-control** = AQC intervention and control indicator; **d-in-d** = pre/post and AQC interaction term; **age**= vector of age categories; **age*sex** = age and sex interaction term; **risk_score** = scalar risk score; **qtr** = vector of quarter time variable; **qtr*AQC** = time and AQC interaction term; **propensity_score** = vector of plan fixed effects where the standard errors are clustered at business entity (practice) level.

	Coefficient*	95% CI	p-value
Pre-Post	-0.0022	(-0.0037 to 0.0006)	0.006
AQC-control	-0.0194	(-0.0252 to 0.0135)	< 0.001
Difference in difference	0.0006	(-0.0018 to 0.0030)	0.606
Model adjusted R^2	0.0266		

* The unit of analysis is total ED visits per member per quarter