# Web Appendix for "Pricing Regulation and Imperfect Competition on the Massachusetts Health Insurance Exchange"

Keith M Marzilli Ericson and Amanda Starc

# A.1 Data Appendix

## A.1.1 Insurance Price Data

#### A.1.1.1 Main Data

For November and December 2009, we used a Perl script to acquire price quotes from the Connector website in various age-zipcode-family size cells. We selected ten geographically distributed zipcodes in which to get a detailed set of price quotes.<sup>30</sup> In these selected zipcodes, we downloaded price quotes for each insurance plan for a single individual for each possible age under 65. (We also obtained family price quotes, but do not analyze them in this paper). Then, for every zipcode in Massachusetts, we downloaded price quotes for each plan for a 30 year old individual.

To build the full choice menu each individual faced, we then constructed estimated prices for all plans for all single individuals in November and December 2009. Using the estimated price quotes from the detailed zipcodes, we estimated the following model separately for each month. The price of plan j for age a in geographical region (market) m is given by

$$p_{ajm} = b_j + b_1 p_{30jm} + \sum_{s \in 1, \dots, 7} \left( 1_{a_{s-1}^* > a \ge a_s^*} \right) \cdot (b_{2s} + b_{3s} p_{30jm})$$

where  $1_{a_{s-1}^*>a\geq a_s^*}$  is an indicator for whether the age for the price quote is in a given age category (between cutoffs  $a_{s-1}^*$  and  $a_s^*$ ). This model which allows for a main effect of plan  $(b_j)$ , age category  $(b_{2s})$  and level of age-30 price, as well as interactions between age category and age-30 price  $(b_{3s})$ . The model has an  $R^2$  of 0.978 on the detailed premium data. Taking the estimated coefficients from this model, we predict  $\hat{p}_{ajm}$  for all j, m, a where we have age 30-price quotes.

#### A.1.1.2 Robustness Checks

Our main analyses use only the November-December data, as we did not download any price quotes from the website prior to November 2009. However, for robustness checks and additional power for some reduced form analyses, we have explored extending the data back until July 2009. Based on the Connector transaction data, we determined that the set of plans j offered was constant between July 2009 and December 2009: no plans appeared or disappeared during this period. The Connector transaction data contains the list prices paid for the plans that individuals actually chose. On their own, these observations are not rich enough to construct the full menu of choices each individual faced. Nonetheless, we can

 $<sup>^{30}</sup>$ These zipcodes are 01020, 01240, 01604, 01824, 01923, 02124, 02130, 02360, 02459, 02474, 02601.

roughly approximate the plan prices that individuals faced between July 2009 and Oct 2009 using our Nov-Dec 2009 prices quotes. For each month t between July and October, we estimated the following equation based on the observed premiums paid by individual i for each plan (among new enrollees for each plan) and the November 2009 price of that plan as follows:

$$p_{i,ajmt} = b_0 + b_1 p_{ajm,Nov2009} + b_{insurer} + b_m$$

where  $b_{insurer}$  and  $b_m$  are month-specific insurer and geographic region effects. We then use this model to predict the prices  $\hat{p}_{ajmt}$  of each plan in each cell. Because we estimate a geographic region fixed effect  $b_m$ , we exclude geographic regions that have fewer than 10 zipcodes from the July-Oct. choice menu construction (and hence from analyses using the July-Oct. data). Our results do not change when the consumer demand model is estimated on this expanded July-Dec. 2009 dataset.

## A.1.2 Massachusetts HIE Transaction Data

We merged the insurance price quotes with our individual level transactions (an add record, cancel record, and payment records). Our main data sample (Nov.-Dec. 2009) focuses on people who enrolled in the exchange for the first time during this time ("add" transactions). We keep observations where the individual is 27-64 years old (inclusive), and who chose individual (not family) plans.

# A.2 Reduced-Form Evidence on Response to Price

This section examines how total spending on health insurance responds to a price index, using a reduced form model. One way to summarize the response to price is the insurance spending elasticity, which relates total premiums paid to an index of the list prices individuals face. We summarize the effect of a price change on insurance spending using the following model:

$$\ln y_i = \eta \ln (p_i) + \gamma \omega_i,$$

where  $y_i$  is the total insurance premiums paid by individual i (given the actual prices),  $p_i$  is a price index for a representative bundle of plans, and  $\omega_i$  is a vector of individual characteristics. The insurance spending elasticity is given by  $\eta$  and says that if the price index rises by 1%, the total spending rises by  $\eta$ %.<sup>31</sup> If  $\eta$  < 1, individuals respond to higher prices by reducing their spending on insurance, while if choice of insurance plan stayed the same, then  $\eta = 1$ .

In this context, the percentage price increase at each threshold varies among plans and insurers.<sup>32</sup> We therefore create a price index, in which each plan is assigned a weight. Because the plan menu varies by geographic region, we create geographic-specific weights: a plan's

 $<sup>^{31}</sup>$ Of course, while identifying  $\eta$  is a valuable way of summarizing the data that can facilitate out-of-context prediction, individuals do not in fact face a continuous choice of dollars spent on health insurance; the discrete choice individuals actually face is modeled in Section 3.

<sup>&</sup>lt;sup>32</sup>By contrast, a change in tax deduction for employer-sponsored health insurance (as in Gruber and Washington 2005) would lead to the same percentage change in price for all the plans, eliminating the need to construct a price index.

weight is the fraction of people in a geographic region who chose that plan, averaged over July to December 2009.<sup>33</sup> Column 1 of Table A.1 shows how the price index jumps at each age threshold. It presents the results of the following regression:

$$\ln p_i = G\left(a\right) + \sum_{s \in 1....7} 1_{a \ge a_s^*} \pi_s + \gamma \omega_i,\tag{1}$$

where G(a) is a linear spline in age and  $\omega_i$  includes gender, month of enrollment, and indicators for geographic region.<sup>34</sup> The coefficients  $\pi_s$  multiply indicator variables for whether age is greater than or equal to each of the age thresholds (each value of  $a_s^*$ ) used for pricing. Each value of  $\pi_s$  shows how the price index jumps at the threshold  $a_s^*$ : for instance, we see that the price index increases by 20.4 log points when an individual turns 50. The jumps in prices are relatively small at age 30 and 35 but are more substantial at older ages.<sup>35</sup>

Next, we examine how total spending on premiums changes at each age threshold shown in Column 2 of Table A.1. It presents the results of the regression

$$\ln y_i = G\left(a\right) + \sum_{s \in 1....7} 1_{a \ge a_s^*} \kappa_s + \gamma \omega_i \tag{2}$$

where G(a),  $\omega_i$ , and the age category indicators are the same as in Equation 1. The values of  $\kappa_s$  from this regression show how spending on premiums jumps at each age discontinuity. Thus, we see spending on premiums jump 19.2 log points at age 55, controlling for linear age trends above and below 55, along with other variables.

Comparing the percentage increase in spending ( $\kappa_s$  in Column 2) to the percentage increase in the price index ( $\pi_s$  from Column 1), we see that the increase in spending is slightly less than the increase in prices for all age thresholds except age 55. The bottom panel of Table A.1 shows the results of an instrumental variable regression that instruments for price index by age-category, controlling for an age spline. (Linear age splines have knots at each age threshold. Additional controls include indicators for month, geographic region, and gender.) Column 1 is thus the first-stage of this IV regression, and the F-statistic for excluded instruments (age-discontinuities) is a substantial 2823. The resulting estimate of  $\eta$  is 0.962 (s.e. 0.176), indicating that a 10% increase in the price index leads to a 9.6% increase in total spending in this population - a relatively limited response by individuals.

This nonstructural approach shows that there is little response in individual choice to price increases: the increase in spending is approximately equal to the increase in prices. Yet Section 3 shows that these results do *not* imply that individuals are not sensitive to price. Rather, individuals are already gravitating to the least generous tier and cheapest plans available. Thus, despite the wide range of plan generosities available in the exchange,

 $<sup>^{33}</sup>$ To construct a reasonable price index, we exclude geographic regions that had fewer than 10 zipcodes, as well as geographic regions that had fewer than four insurers.

<sup>&</sup>lt;sup>34</sup>Because we know the pricing model, gender and the linear age spline do not predict prices; they are included for comparability with later regressions.

<sup>&</sup>lt;sup>35</sup>The measured increases in the price index at the age thresholds (and their relative magnitudes) vary depending on how the price index is constructed. We gave plans weights based on popularity in a geographical region. Ideally, we would assign age-group specific weights (to create a Laspeyres or Paasche index). However, the sparsity of the data makes this an unappealing route. The construction of the price index has no bearing on how we measure the change in spending at each threshold (but is relevant for estimating  $\eta$ ).

individuals do not have much latitude to respond to a price increase. They are simply unable to substitute to cheaper plans. Thus, these results highlight the importance of context in determining the effect of policy changes (e.g., altering the tax exclusion for employer-provided health insurance) and motivate our structural model of consumer preferences in Section 3.

# A.3 Appendix Figures and Tables

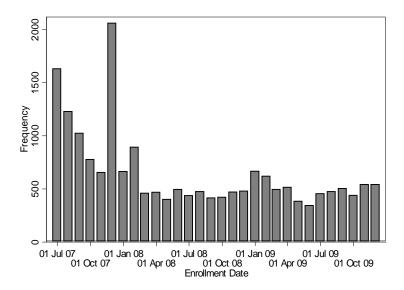


Figure A.1: Distribution of Initial Enrollment of Single Individuals.

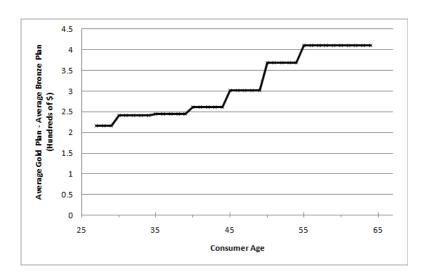


Figure A.2: Marginal Cost of Average Gold Plan versus Average Bronze Plan, By Age. Unit of observation is a plan-zipcode in Nov. 2009.

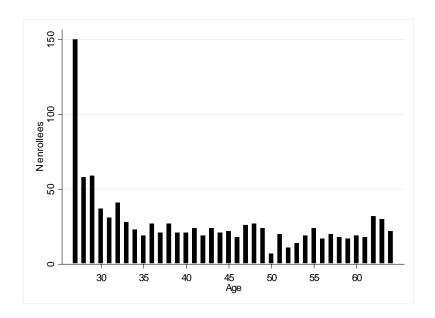


Figure A.3: Number of New Enrollees, By Age. Sample: Nov.-Dec 2009.

Table A.1: Price and Spending Response to Age Discontinuities

|                   | ln(price index) | ln(premiums paid) |
|-------------------|-----------------|-------------------|
| Indicators:       |                 |                   |
| Above 30          | 0.0224***       | -0.0438           |
|                   | (0.00172)       | (0.0323)          |
| Above 35          | 0.0790***       | 0.0442            |
|                   | (0.00199)       | (0.0411)          |
| Above 40          | 0.150***        | 0.147***          |
|                   | (0.00217)       | (0.0447)          |
| Above 45          | 0.106***        | 0.0138            |
|                   | (0.00189)       | (0.0440)          |
| Above 50          | 0.204***        | 0.207***          |
|                   | (0.00201)       | (0.0502)          |
| Above 55          | 0.128***        | 0.192***          |
|                   | (0.00232)       | (0.0462)          |
| Linear Age Spline | Yes             | Yes               |
| Basic Controls    | Yes             | Yes               |
| N Persons         | 2,616           | 2,616             |
| $R^2$             | 0.998           | 0.572             |

## IV-Stage 1 from Column 1

 $\begin{array}{c} & \text{ln(premiums paid)} \\ \text{ln(price index)} & 0.962 \\ & (0.176) \\ \text{Linear Age Spline} & \text{Yes} \\ \text{Basic Controls} & \text{Yes} \\ R^2 & 0.569 \\ \end{array}$ 

Sample: July-Dec. 2009 extended data sample. Note: Heteroskedasticity robust standard errors in parentheses. Age spline consists of piecewise linear age controls within each age group. Controls include indicators for month of enrollment, indicators for geographic market, and gender. IV results from two-stage least squares. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.2: Zipcode Characteristics of Enrollees By Age

|                  |            | Fraction   |            | Zip Cod     | e Income: | Count               |
|------------------|------------|------------|------------|-------------|-----------|---------------------|
|                  | Unemployed | Married    | White      | Mean        | Median    | (in 1 year age bin) |
|                  |            |            | Omitted Ca | tegory:Age  | 27-29     |                     |
| Above 30         | 0.00295    | -0.00778   | -0.0411    | -2,739      | -1,499    | -9.107              |
|                  | (0.00318)  | (0.0260)   | (0.0291)   | (2,392)     | (3,992)   | (24.51)             |
| Above 35         | 0.00387    | 0.0157     | -0.0352    | -2,947      | 822.7     | 21.98               |
|                  | (0.00402)  | (0.0246)   | (0.0340)   | (2,620)     | (3,722)   | (18.89)             |
| Above 40         | -0.00753** | 0.0319     | 0.0265     | 894.0       | 6,051     | 4.412               |
|                  | (0.00323)  | (0.0253)   | (0.0272)   | (2,280)     | (4,039)   | (6.624)             |
| Above 45         | 0.00153    | -0.000514  | -0.0354    | -902.8      | 1,518     | -8.676              |
|                  | (0.00319)  | (0.0242)   | (0.0357)   | (2,307)     | (4,006)   | (7.041)             |
| Above 50         | -0.00202   | -0.0150    | -0.0284    | -761.4      | 219.0     | -26.29**            |
|                  | (0.00422)  | (0.0257)   | (0.0291)   | (2,471)     | (4,483)   | (9.906)             |
| Above 55         | -0.00696** | 0.0334     | 0.0474*    | 3,730**     | 6,887*    | -9.959              |
|                  | (0.00319)  | (0.0212)   | (0.0250)   | (1,827)     | (3,524)   | (8.320)             |
| Age              | -0.00544   | -0.0188    | 0.00433    | 2,252       | -819.5    | -90.25*             |
| 9                | (0.00675)  | (0.0480)   | (0.0552)   | (4,564)     | (8,011)   | (51.01)             |
| $Age^2$          | 0.000118   | 0.000417   | 0.000118   | -34.93      | 13.87     | 1.847*              |
| O                | (0.000154) | (0.00109)  | (0.00127)  | (103.1)     | (184.2)   | (1.050)             |
| $\mathrm{Age^3}$ | -7.90e-07  | -2.95e-06  | -1.81e-06  | $0.147^{'}$ | -0.131    | -0.0121*            |
| S                | (1.10e-06) | (7.77e-06) | (9.10e-06) | (0.734)     | (1.326)   | (0.00697)           |
| Constant         | 0.105      | 0.771      | 0.670      | -10,126     | 70,991    | 1,433*              |
|                  | (0.0933)   | (0.667)    | (0.759)    | (63,563)    | (109,728) | (772.0)             |
| Observations     | 1,052      | 1,052      | 1,052      | 1,052       | 1,052     | 38                  |
| $R^2$            | 0.019      | 0.050      | 0.032      | 0.007       | 0.012     | 0.767               |

Note: Heteroskedasticity robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Zipcode demographics are taken from the 2000 Census.

Table A.3: Price Sensitivity by Age in Conditional Logit Model, Bandwidth Checks

|                               | 27-34   | 30-39   | 36-44   | 40-49   | 46-54   | 50+   |
|-------------------------------|---|---|---|---|---|---|
| Premium (in \$100s)           | -3.632***<br>(0.740)                              | -2.575***<br>(0.720)                              | -2.005***<br>(0.773)                              | -2.700***<br>(0.735)                              | -1.738**<br>(0.767)                               | -1.375<br>(0)                                     |
| Observations<br>Fixed Effects | 2944<br>Plan<br>Tier*Age<br>Tier*Age <sup>2</sup> | 3098<br>Plan<br>Tier*Age<br>Tier*Age <sup>2</sup> | 2615<br>Plan<br>Tier*Age<br>Tier*Age <sup>2</sup> | 2575<br>Plan<br>Tier*Age<br>Tier*Age <sup>2</sup> | 2194<br>Plan<br>Tier*Age<br>Tier*Age <sup>2</sup> | 1701<br>Plan<br>Tier*Age<br>Tier*Age <sup>2</sup> |

Note: Heteroskedasticity robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Panels B and C include plan fixed effects, and tier effects interacted with age trends (both linear and quadratic terms).

Table A.4: Direct Purchase of Insurance in Mass. by Age, American Community Survey

| Age Group | N Direct Purchase | Fraction of Sample | Fraction of sample,<br>normalized by years in age bin |
|-----------|-------------------|--------------------|---|
| 27-29*    | 47,432            | 0.09               | 0.030   |
| 30-34     | 67,502            | 0.13               | 0.026   |
| 35-39     | 66,054            | 0.13               | 0.025   |
| 40-44     | 75,878            | 0.15               | 0.029   |
| 45-49     | 77,910            | 0.15               | 0.030   |
| 50-54     | 69,576            | 0.13               | 0.027   |
| 55-59     | 60,652            | 0.12               | 0.023   |
| 60-64     | 54,823            | 0.11               | 0.021   |

Data taken from American Community Survey, 3-year estimates, 2008-2010, Massachusetts only. Direct purchase variable is "HINS2". \*Note smaller bin size in first row.

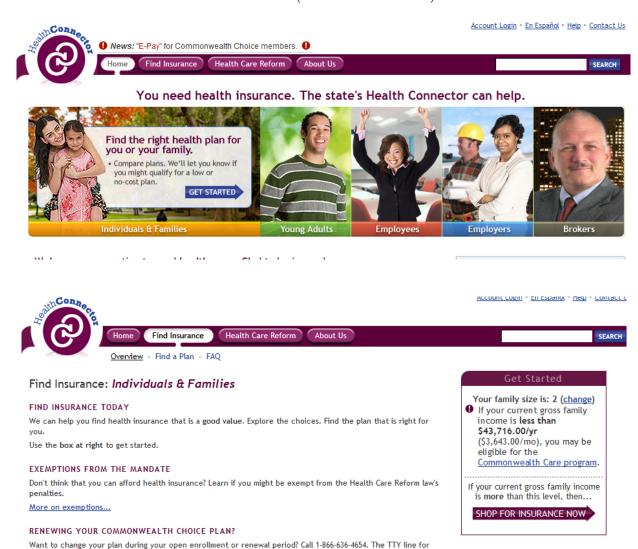
Table A.5: Price Sensitivity by Age in Conditional Logit Model, Bronze Only

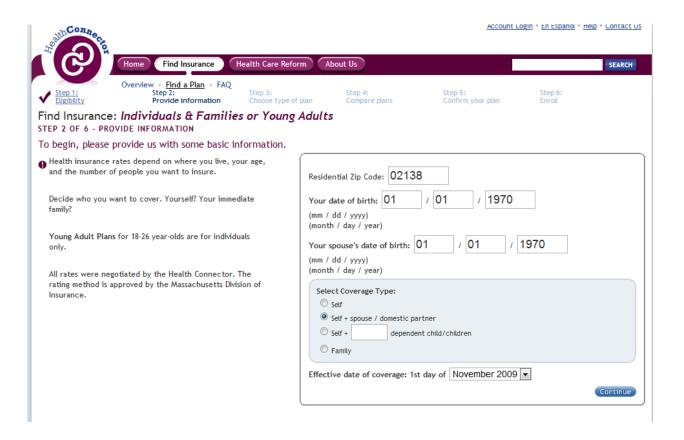
|               | Panel A      | : Basic Cond   | ditional Logit | ts (Full Sam  | ole)         |              |
|---------------|--------------|----------------|----------------|---------------|--------------|--------------|
|               | (:           | 1)             | ('.            | 2)            | ;)           | 3)           |
| Premium       | -0.31        | 3***           | -1.71          | 9***          | -1.74        | 9***         |
| (in \$100s)   | (0.1         | 109)           | (0.2)          | 242)          | (0.2)        | 254)         |
| Premium*age   |              |                | 0.027          | 71***         | 0.00         | )312         |
|               |              |                | (0.00)         | 0419)         | (0.00)       | 0871)        |
| Fixed Effects | Pl           | an             | Pl             | an            | Plan, P      | lan*Age      |
| N Person*Plan | 6,8          | 370            | 6,8            | 370           | 6,8          | 370          |
|               | Pane         | l B: Condition | onal Logits b  | y Age Group   | )            |              |
|               | 27-34        | 30-39          | 36-44          | 40-49         | 46-54        | 50+          |
| Premium       | -4.517***    | -4.428***      | -2.172***      | -1.990***     | -1.832***    | -1.402***    |
| (in \$100s)   | (0.662)      | (0.841)        | (0.697)        | (0.627)       | (0.626)      | (0.326)      |
| N Person*Plan | 2803         | 1175           | 1145           | 1471          | 1235         | 1858         |
| Fixed Effects | Plan         | Plan           | Plan           | Plan          | Plan         | Plan         |
|               | Tier*Age     | Tier*Age       | Tier*Age       | Tier*Age      | Tier*Age     | Tier*Age     |
|               | $Tier*Age^2$ | $Tier*Age^2$   | $Tier*Age^2$   | $Tier*Age^2$  | $Tier*Age^2$ | $Tier*Age^2$ |
|               | Panel C: Co  | onditional Lo  | gits For Cou   | nterfactual I | Exercise     |              |
|               |              | Under $45$     |                |               | Over 45      |              |
| Premium       |              | -3.291***      |                |               | -1.463***    |              |
| (in \$100s)   |              | (0.464)        |                |               | (0.205)      |              |
| Fixed Effects |              | Plan           |                |               | Plan         |              |
|               |              | Tier*Age       |                |               | Tier*Age     |              |
|               |              | $Tier*Age^2$   |                |               | $Tier*Age^2$ |              |
| N Person*Plan |              | 4248           |                |               | 6870         |              |

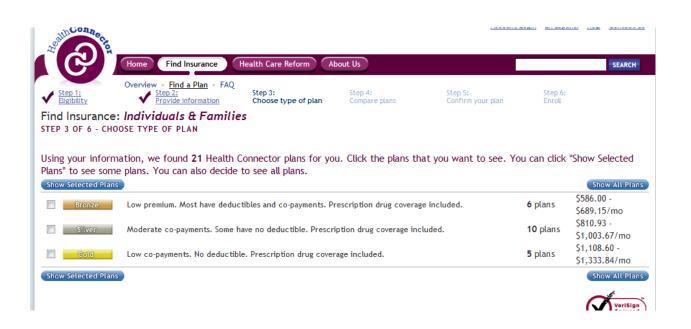
Note: Heteroskedasticity robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Panels B and C include plan fixed effects, and tier effects interacted with age trends (both linear and quadratic terms).

# A.4 Screenshots from the Choice Process in the Exchange

Screenshots from the Massachusetts HIE (circa October 2009) are below:









## Find Insurance: Individuals & Families

STEP 4 OF 6 - COMPARE PLANS (OVERVIEW)

Click "View Plan" to see details. You can also compare up to 3 plans at a time. Check the box next to the plans you want to compare. Then click "Compare Selected Plans."

#### Compare Selected Plans

|             |   |             |   | Co-Pay  | ments 🕖  |                           |  | Doctors        |                            |
|-------------|---|-------------|---|---|--|---------------------------|--|----------------|----------------------------|
| <u>Tier</u> | <u>Plan</u>   | Premium* () | Deductible                                    | Doctor  | <u>RX</u>  | ER                        | Hospital Stay                              | Can<br>See 🚱   | Choose<br>Plan             |
| В           | Fallon Community Health<br>Plan<br>FCHP Direct Care                                 | \$586.00    | \$2,000/\$4,000                               | \$25  | \$15 / \$50 / \$100  | \$200                     | \$500 per<br>admission after<br>deductible | Find<br>Doctor | View<br>Plan               |
| В           | Neighborhood Health Plan NHPThree Select  | \$636.22    | \$2,000/\$4,000                               | \$25  | \$15 after Rx deductible<br>/ 50% co-insurance<br>after Rx deductible /<br>50% co-insurance after<br>Rx deductible | \$100 after<br>deductible | 20% co-insurance<br>after deductible       | Find<br>Doctor | <u>View</u><br><u>Plan</u> |
| В           | Harvard Pilgrim Health Care Harvard Pilgrim Core Coverage 1750                      | \$641.71    | \$1,750/\$3,500                               | \$25 copay up to 3 medical care<br>office visits per individual (or 6<br>per family); next visits are<br>subject to the deductible; then<br>20% co-insurance thereafter | \$15 / 50% co-insurance<br>after Rx deductible /<br>50% co-insurance after<br>Rx deductible                        | \$250                     | 20% co-insurance<br>after deductible       | Find<br>Doctor | <u>View</u><br><u>Plan</u> |
| В           | Fallon Community Health<br>Plan<br>FCHP Select Care                                 | \$676.00    | \$2,000/\$4,000                               | \$25  | \$15 / \$50 / \$100  | \$200                     | \$500 per<br>admission after<br>deductible | Find<br>Doctor | View<br>Plan               |
| В           | Tufts Health Plan Advantage HMO Select 2000 (Limited choice of doctors & hospitals) | \$676.73    | \$2,000/\$4,000                               | \$40  | \$20 after Rx deductible<br>/ \$50 after Rx<br>deductible / \$75 after<br>Rx deductible                            | \$200                     | \$0 after<br>deductible                    | Find<br>Doctor | View<br>Plan               |
| В           | Blue Cross Blue Shield of<br>Massachusetts<br>HMO Blue Basic Value                  | \$689.15    | \$250 per plan<br>year/\$500 per<br>plan year | \$25  | \$15 / 50% co-insurance<br>after Rx deductible /<br>50% co-insurance after<br>Rx deductible                        | \$200                     | 35% co-insurance<br>after deductible       | Find<br>Doctor | View<br>Plan               |
| S           | Tufts Health Plan Advantage HMO Select 750 (Limited choice of doctors &             | \$810.93    | \$750/\$1,500                                 | \$15  | \$10 after Rx deductible<br>/ \$30 after Rx<br>deductible / \$45 after   | \$200                     | \$0 after<br>deductible                    | Find<br>Doctor | <u>View</u><br><u>Plan</u> |

#### Find Insurance: Individuals & Families

STEP 4 OF 6 - COMPARE PLANS (DETAILS)

Here are the details of the plan(s) that you are comparing.

• Note: Premiums as of 9/25/2009 for an effective date of 11/1/2009.

|   | Choose Plan                              |
|---|--|
| CARRIER NAME  | FALLON COMMUNITY HEALTH PLAN             |
| Plan Name   | FCHP Direct Care                         |
| With or Without Pharmacy (Rx)   | With Rx                                  |
| O Connector Plan Tier   | Bronze                                   |
| PLAN DETAILS  | <u>Download Plan Details</u>             |
| PREMIUM   | \$586.00                                 |
| ANNUAL DEDUCTIBLE 1 Per person  | \$2,000                                  |
| Per person  | \$2,000                                  |
| Family total  | \$4,000                                  |
|   |  |
| ANNUAL OUT-OF-POCKET (OOP) MAXIMUM <sup>2</sup>   |  |
| Per person  | \$5,000                                  |
| Per person<br>Family total  |  |
| Per person Family total Costs that count towards OOP maximum  | \$5,000<br>\$10,000                      |
| Per person Family total Costs that count towards 00P maximum Golfice visit: Adult routine physical  | \$5,000<br>\$10,000<br>Yes               |
| Per person Family total Costs that count towards OOP maximum  Golffice visit: Adult routine physical Golffice visit: Routine gynecological (GYN) exam | \$5,000<br>\$10,000<br>Yes<br>Yes        |
| Per person Family total Costs that count towards OOP maximum  | \$5,000<br>\$10,000<br>Yes<br>Yes<br>Yes |
| Per person Family total Costs that count towards OOP maximum  Diffice visit: Adult routine physical Diffice visit: Routine gynecological (GYN) exam   | \$5,000<br>\$10,000<br>Yes<br>Yes        |