Online Appendix:
Price-Linked Subsidies and Imperfect Competition in Health Insurance[*]

Sonia Jaffe and Mark Shepard

November 8, 2019

A   Theory

A.1   Multi-Plan Insurers (ACA Case)

In this section, we show how the basic theory of price-linked subsidies' effect on markups (see Section I) carries over to the multi-plan setting of the ACA. In the ACA insurers may offer plans in each of multiple tiers – bronze, silver, gold, and platinum. Subsidies are set equal to the price of the second-cheapest silver plan minus a pre-specified "affordable amount."[1] In general, the fact that insurers are providing additional plans provides a greater incentive for an insurer to increase the price of its silver plan because the higher subsidy increases demand for the insurer's non-silver plans as well – again by inducing more customers to enter the market.[2]

Suppose each firm j offers plans in tiers $l = \{(B)ronze, (S)ilver, (G)old, (P)latinum\}$. For notational simplicity, assuming no adverse selection (or perfect risk adjustment). The insurer maximizes profits:

$$\max_{P_{j1}...P_{jL}} \sum_l (P_{jl} - c_{jl}) Q_{jl} (P^{cons}, M),$$

where $P_{jl}^{cons} = P_{jl} - S(P)$ with $S(P) = P_{2nd,S} - \text{AffAmt}$. Following the same steps as in the text, the first-order condition for the silver plan is:

$$\frac{\partial \pi_j}{\partial P_{jS}} = Q_{jS}(\cdot) + \sum_l (P_{jl} - c_{jl}) \frac{dQ_{jl}}{dP_{jS}} = 0.$$

---

[1]The subsidy applies equally to all plans (though no premium can go below $0), ensuring that at least two silver plans (and likely some bronze plans) cost low-income consumers less than the affordable amount.

[2]If an insurer does not offer a bronze plan *and* consumers on the margin of uninsurance mostly pick bronze plans, that would mitigate the distortion. Empirically, however, nearly all participating insurers do offer at least one bronze plan.

The markup with fixed subsidies is:

$$Mkup_{j\mathrm{S}}^{\mathrm{F}} = \frac{1}{\eta_{j\mathrm{S}}} + \frac{1}{-\frac{\partial Q_{j\mathrm{S}}}{\partial P_{j\mathrm{S}}^{cons}}} \sum_{l \neq silv} (P_{jl} - c_{jl}) \frac{\partial Q_{jl}}{\partial P_{j\mathrm{S}}^{cons}}$$

The second term reflects a standard effect for multi-product firms: when the insurer raises the price of one plan, it captures revenue from consumers who switch to its other plans. However, with price-linked subsidies, the markup for the subsidy-pivotal plan is:

$$Mkup_{j\mathrm{S}}^{\mathrm{PLink}} = \frac{1}{\eta_{j\mathrm{S}} - \eta_{j\mathrm{S},M}} + \frac{\sum_{l \neq S} (P_{jl} - c_{jl}) \left(\frac{\partial Q_{jl}}{\partial P_{j\mathrm{S}}^{cons}} + \overbrace{\frac{\partial Q_{jl}}{\partial M}}^{\text{Additional Distortion}} \right)}{-\frac{\partial Q_{j\mathrm{S}}}{\partial P_{j\mathrm{S}}^{cons}} - \underbrace{\frac{\partial Q_{j\mathrm{S}}}{\partial M}}_{\text{Additional Distortion}}}.$$

The fact that other plans offered by the firms also gain some of the consumers driven into the market by the additional subsidy generates an additional distortion.

A.2 Adverse Selection

Our basic theory in Section I abstracted from adverse selection (beyond what the exchange's risk adjustment accounted for) for expositional simplicity. In this section, we show that the basic logic of the distortion of price-linked subsidies generalizes to a model with risk selection. With risk selection, a plan's average costs, $\bar{c}_j(P^{cons}, M)$, depend on the set of consumers who select a plan, which is affected by premiums and the mandate penalty. The exchange uses a risk adjustment transfer, $\bar{\phi}_j(P^{cons}, M)$, to compensate plans based on the measured sickness of its enrollees (which also varies with prices). The insurer's net (risk-adjusted) costs equal:

$$c_j^{Net}(P) = \bar{c}_j(P^{cons}, M) - \bar{\phi}_j(P^{cons}, M).$$

The insurer profit function is:

$$\pi_j = \left(P_j - c_j^{Net} \right) \cdot Q_j \left(P^{cons}, M \right).$$

It's first-order condition is

$$\frac{d\pi_j}{dP_j} = \left(1 - \frac{dc_j^{Net}}{dP_j} \right) Q_j \left(P^{cons}, M \right) + \left(P_j - c_j^{Net} \right) \cdot \frac{dQ_j}{dP_j} = 0.$$

Just like the total derivative of demand, the total derivative of cost, $\frac{dc_j^{Net}}{dP_j}$, differs with the

subsidy rule. With fixed subsidies

$$\frac{dc_j^{Net}}{dP_j} = \frac{\partial c_j^{Net}}{\partial P_j^{cons}}.$$

With price-linked subsidies

$$\frac{dc_j^{Net}}{dP_j} = \frac{\partial c_j^{Net}}{\partial P_j^{cons}} + \frac{\partial c_j^{Net}}{\partial M}.$$

The respective markups are

$$Mkup_j^{\mathrm{F}} = \frac{1}{\eta_j} \left(1 - \frac{\partial c^{Net}}{\partial P_j^{cons}} \right)$$

$$Mkup_{\underline{j}}^{\mathrm{PLink}} = \frac{1}{\eta_{\underline{j}} - \eta_{\underline{j},M}} \left(1 - \frac{\partial c^{Net}}{\partial P_j^{cons}} - \frac{\partial c^{Net}}{\partial M} \right).$$

where $\eta_j \equiv -\frac{1}{Q_j}\frac{\partial Q_j}{\partial P_j^{cons}}$ is the own-price semi-elasticity of demand and $\eta_{\underline{j},M} \equiv \frac{1}{Q_j}\frac{\partial Q_j}{\partial M}$ is the semi-elasticity of demand for $\underline{j}$ with respect to the mandate penalty. The difference is

$$Mkup_{\underline{j}}^{\mathrm{PLink}} - Mkup_{\underline{j}}^{\mathrm{F}} = \frac{\eta_{\underline{j},M}}{\eta_{\underline{j}}\left(\eta_{\underline{j}} - \eta_{\underline{j},M}\right)} \left(1 - \frac{\partial c^{Net}}{\partial P_j^{cons}} \right) - \frac{\partial c^{Net}}{\partial M}\frac{1}{\eta_{\underline{j}} - \eta_{\underline{j},M}}. \tag{A.1}$$

Adverse selection implies lower markups under either fixed or price-linked subsidies, since $\left(1 - \frac{\partial c^{Net}}{\partial P_j} \right) < 1$. This decrease in markups has the effect of decreasing the difference in markups between the subsidy policies. However, with adverse selection, a higher subsidy brings healthier people into the market, decreasing costs ($\frac{\partial c^{Net}}{\partial M} < 0$), so the second term is positive, raising the difference in markups. The change as a fraction of the markup under fixed subsidies

$$\frac{Mkup_{\underline{j}}^{\mathrm{PLink}} - Mkup_{\underline{j}}^{\mathrm{F}}}{Mkup_{\underline{j}}^{\mathrm{F}}} = \frac{\eta_{\underline{j},M}}{\left(\eta_{\underline{j}} - \eta_{\underline{j},M}\right)} - \frac{\frac{\partial c^{Net}}{\partial M}}{1 - \frac{\partial c^{Net}}{\partial P_j^{cons}}}\frac{\eta_{\underline{j}}}{\eta_{\underline{j}} - \eta_{\underline{j},M}}$$

is larger with adverse selection because of the second term (compared to Equation (4) in the main text).

A.3 Alternate Policy: Price-Linked Subsidies and Mandate Penalty

Our model suggests an alternative subsidy structure to eliminate the price distortion while still guaranteeing affordability of post-subsidy premiums, as with price-linked subsidies. Specifically, regulators could set a base mandate penalty M_0 and then *apply the subsidy*

to the mandate penalty (in addition to the insurance plans) so that:

$$M\left(P\right) = M_0 - S\left(P\right).$$

The key feature of this policy is that $\frac{\partial M}{\partial P_j} = -\frac{\partial S}{\partial P_j}$, so the "Price-Linking" term in Equation (3) equals zero. As a result, the subsidy policy does not diminish the effective slope of the demand curve. The government could set M_0 so that in expectation, $M(P)$ would equal the penalty under the current system, but the actual mandate penalty would depend on market prices.

Intuitively, this works because it holds fixed $S+M$, the *net* public incentive for consumers to buy insurance. (Fixed subsidies do the same by holding *both* S and M fixed with prices.) Since plans' pricing cannot impact this net incentive for insurance, the distortion of price-linking goes away. However, this policy loses most of the benefits of price-linked subsidies discussed in Section I.B. The net incentive to buy insurance is not linked to the externality of uninsurance (since it does not vary with prices). Also, while the policy removes the costly variation in consumer premiums that fixed subsidies have, but the variation is transferred to the mandate penalty, which now varies with prices; the price-risk is transferred from consumers who buy insurance to those who do not. Of course, this policy only eliminates the distortion if the net incentive, $S + M$, is what matters for insurance demand, not the level of S and M individually. If some consumers were unaware of the penalty or could avoid paying it (e.g., by applying for a religious or hardship exemption), this assumption would not hold perfectly. In our empirical work, therefore, we focus on comparing fixed and price-linked subsidies, rather than this alternate policy.

A.4   Regulator's Objective Function: Public Surplus

We base our welfare analysis on a consumer surplus standard (common in antitrust analysis), adjusted to subtract net government spending (subsidy cost net of mandate revenue) and a social externality of uninsurance. Together, these form a regulatory objective that we refer to as "public surplus." The expected public surplus from individual i equals

$$PS^i = \underbrace{CS^i\left(P^{cons}, M\right)}_{\text{Cons. Surplus}} + \underbrace{\left(1 - D_0^i\left(P^{cons}, M\right)\right) E^i}_{\text{Avoided Externality}} - \underbrace{((1 - D_0^i(P^{cons}, M))S - D_0^i(P^{cons}, M)M)}_{\text{Net Govt. Cost}},$$

where D_0^i is the probability the individual does not chose insurance and E^i is the externality of that individual being uninsured, which is avoided when they purchase insurance. In practice, we treat this social disutility as a free parameter that we calibrate to make the observed subsidy levels (or post-subsidy consumer premiums) optimal.

We think this public surplus standard is consistent with the public health care reform debate, which put little weight on insurer profits – indeed, policymakers seemed eager to constrain insurer profits with policies like medical loss ratio limits – but we also consider an alternative where the regulator values insurer profits, maximizing $PS + \sum_j \pi_j$.

4

Incorporating Risk Aversion

Our baseline simulations calculate consumer welfare assuming risk-neutral consumers, but we also consider a modification of public surplus that captures the benefit of stabilizing post-subsidy consumer prices, transferring the risk of cost shocks to the government.

Let consumers have concave utility over non-health insurance consumption, $u(\cdot)$, with a (local) coefficient of relative risk aversion $\gamma = -\frac{u''(y)y}{u'(y)}$. If Y is income and P is the premium for insurance, then the change in utility from a premium difference of ΔP is

$$
\begin{aligned}
u(Y - P - \Delta P) - u(Y - P) &\approx -u'(Y - P)\Delta P + u''(Y - P)\frac{(\Delta P)^2}{2} \\
&= -u'(Y - P)\left(\Delta P + \frac{\gamma}{Y - P}\frac{(\Delta P)^2}{2}\right) \\
&= -\Delta P - \frac{\gamma}{Y - P}\frac{(\Delta P)^2}{2}.
\end{aligned}
\tag{A.2}
$$

The last line follows from the fact that in order to combine consumer surplus and government spending, we implicitly assume that prior to any cost shocks, the transfer system was set "right" so the consumer's marginal utility of a dollar ($u'(Y - P_{j*})$) equals the marginal value of a dollar to the government (which is normalized to one). When weighted by demand shares, the first term is the effect on risk neutral consumers which is included in our baseline calculation. The second term is the additional cost from risk aversion, which we add to baseline public surplus to capture risk protection.

A.5 Optimality of Fixed Subsidies without Uncertainty

To show that fixed subsidies are optimal, we model there being a total subsidy, S that is a mixture between a fixed subsidy $\hat{S}$ and a subsidy that depends on the cheapest plan's price with weight α:

$$
S = \hat{S} + \alpha P_{\underline{j}},
$$

where firm $\underline{j}$ is the subsidy-pivotal plan. We show that under the conditions Section I.B, as long as the level of $\hat{S}$ is set optimally, the derivative of welfare with respect to α is negative, implying that it is optimal to not link subsidies to price. (This actually suggests that the optimal subsidy would be negatively linked to prices, but that is not a policy we have seen proposed.)

Consumers only care about the total subsidy, which is also what determines government expenditures. Therefore, α and $\hat{S}$ affect welfare only via their effects on the total subsidy

and their effects on prices.[3] So,

$$\frac{dW}{d\hat{S}} = \frac{\partial W}{\partial S} + \sum_j \frac{\partial W}{\partial P_j}\frac{\partial P_j}{\partial \hat{S}}$$

$$\frac{dW}{d\alpha} = \frac{\partial W}{\partial S} \cdot P_{\underline{j}} + \sum_j \frac{\partial W}{\partial P_j}\frac{\partial P_j}{\partial \alpha}$$

If $\hat{S}$ is set optimally, then the first line equals zero. Substituting the resulting expression for $\frac{\partial W}{\partial S}$ into the second line, it becomes

$$\frac{dW}{d\alpha} = \sum_j \frac{\partial W}{\partial P_j}\left(\frac{\partial P_j}{\partial \alpha} - \frac{\partial P_j}{\partial \hat{S}} \cdot P_{\underline{j}}\right).$$

Condition (iii) is that $\frac{\partial W}{\partial P_j}$ is negative for all j, so $\frac{\partial W}{\partial \alpha}$ is negative as long as the price changes $\left(\frac{\partial P_j}{\partial \alpha} - \frac{\partial P_j}{\partial \hat{S}} \cdot P_{\underline{j}}\right)$ are positive.

For price changes, we turn to the firms' first-order conditions, which must continue to hold as α or $\hat{S}$ changes. If π_i is each firm's profit function, the effect of $\hat{S}$ and α on prices follow

$$\left[\frac{\partial^2 \pi_i}{\partial P_i \partial P_j}\right] \cdot \frac{\partial P}{\partial \hat{S}} + \frac{\partial^2 \pi_i}{\partial P_i \partial \hat{S}} = 0$$

$$\left[\frac{\partial^2 \pi_i}{\partial P_i \partial P_j}\right] \cdot \frac{\partial P}{\partial \alpha} + \frac{\partial^2 \pi_i}{\partial P_i \partial \alpha} = 0$$

so

$$\frac{\partial P}{\partial \alpha} - P_{\underline{j}} \cdot \frac{\partial P}{\partial \hat{S}} = -\left[\frac{\partial^2 \pi_i}{\partial P_i \partial P_j}\right]^{-1}\left(\frac{\partial^2 \pi_i}{\partial P_i \partial \alpha} - P_{\underline{j}}\frac{\partial^2 \pi_i}{\partial P_i \partial \hat{S}}\right). \tag{A.3}$$

For all firms except $\underline{j}$, both $\hat{S}$ and α only enter the profit function via S so

$$\frac{\partial^2 \pi_i}{\partial P_i \partial \alpha} = P_{\underline{j}}\frac{\partial^2 \pi_i}{\partial P_j \partial S} = P_{\underline{j}}\frac{\partial^2 \pi_i}{\partial P_j \partial \hat{S}},$$

making the term in parentheses in Equation (A.3) zero. For firm $\underline{j}$,

$$\frac{\partial \pi_{\underline{j}}}{\partial P_{\underline{j}}} = \left(1 - \frac{\partial c_{\underline{j}}^{Net}}{\partial P_{\underline{j}}}\right) + \frac{\partial Q_{\underline{j}}}{\partial P_{\underline{j}}}(P_{\underline{j}} - c_{\underline{j}}^{Net}) + \alpha\left(\frac{\partial Q_{\underline{j}}}{\partial \hat{S}}(P_{\underline{j}} - c_{\underline{j}}^{Net}) - Q_{\underline{j}}\frac{\partial c_{\underline{j}}^{Net}}{\partial \hat{S}}\right)$$

$$\frac{\partial^2 \pi_{\underline{j}}}{\partial P_{\underline{j}} \partial \alpha} - P_{\underline{j}}\frac{\partial^2 \pi_{\underline{j}}}{\partial P_{\underline{j}} \partial \hat{S}} = \left(\frac{\partial Q_{\underline{j}}}{\partial \hat{S}}(P_{\underline{j}} - c_{\underline{j}}^{Net}) - Q_{\underline{j}}\frac{\partial c_{\underline{j}}^{Net}}{\partial \hat{S}}\right).$$

---

[3]Even if profits enter welfare, ex-post these only depend on the total subsidy, though ex-ante, α and $\hat{S}$ obviously create different incentives.

This is equal to $= P_{\underline{j}} \frac{\partial \pi_{\underline{j}}}{\partial \hat{S}}$, which Condition (i) requires to be positive.

Condition (ii) ensures that the effect on other firm's prices when the pivotal plan has a marginal cost increase,

$$\frac{\partial P_i}{\partial c_{\underline{j}}^{Net}} = -\left[\frac{\partial^2 \pi_i}{\partial p_i \partial p_j}\right]_{j,\underline{j}}^{-1},$$

is positive. Since all the terms in the $\underline{j}^{th}$ column of $\left[\frac{\partial^2 \pi_i}{\partial p_i \partial p_j}\right]^{-1}$ are negative and $\frac{\partial^2 \pi_{\underline{j}}}{\partial P_{\underline{j}} \partial \alpha} - P_{\underline{j}} \frac{\partial^2 \pi_{\underline{j}}}{\partial P_{\underline{j}} \partial \hat{S}}$ are positive, $\frac{\partial P_i}{\partial \alpha} - P_{\underline{j}} \cdot \frac{\partial P_i}{\partial \hat{S}}$ is positive for all j, implying that $\frac{dW}{d\alpha}$ is negative.

Uncertainty

The above logic breaks down under uncertainty. With uncertainty, the first-order condition for the fixed component of the subsidy only holds in expectation:

$$\frac{dW}{d\hat{S}} = E\left[\frac{\partial W}{\partial S} + \sum_j \frac{\partial P_j}{\partial \hat{S}} \frac{\partial W}{\partial P_j}\right] = 0.$$

This means that

$$E[P_{\underline{j}}] \cdot E\left[\frac{\partial W}{\partial S} + \sum_j \frac{\partial P_j}{\partial \hat{S}} \frac{\partial W}{\partial P_j}\right] = 0$$

$$cov\left(P_{\underline{j}}, \frac{\partial W}{\partial S} + \sum_j \frac{\partial P_j}{\partial \hat{S}} \frac{\partial W}{\partial P_j}\right) - E\left[P_{\underline{j}}\left(\frac{\partial W}{\partial S} + \sum_j \frac{\partial P_j}{\partial \hat{S}} \frac{\partial W}{\partial P_j}\right)\right] = 0.$$

So the effect of α is

$$\frac{dW}{d\alpha} = E\left[\frac{\partial W}{\partial S} \cdot P_{\underline{j}} + \sum_j \frac{\partial P_j}{\partial \alpha} \frac{\partial W}{\partial P_j}\right]$$

$$= cov\left(P_{\underline{j}}, \frac{\partial W}{\partial S} + \sum_j \frac{\partial P_j}{\partial \hat{S}} \frac{\partial W}{\partial P_j}\right) + E\left[\sum_j \frac{\partial W}{\partial P_j}\left(\frac{\partial P_j}{\partial \alpha} - P_{\underline{j}} \frac{\partial P_j}{\partial \hat{S}}\right)\right].$$

Under uncertainty, price-linking can only have a positive effect on welfare if the covariance between the subsidy-pivotal price and the marginal effect of the subsidy on welfare is sufficiently positive to outweigh the negative effect of the higher prices.

B Data and Reduced Form Estimation Details

B.1 ACS Data Construction

We use the American Community Survey (ACS) to estimate the size of the CommCare-eligible population that chooses uninsurance and to generate a micro dataset of these indi-

viduals for demand estimation and simulations. For each CommCare year, we pool observations for the two ACS calendar years in which it occurred. For instance, for CommCare year 2009 (which ran from July 2008 to June 2009), we pool observations from ACS 2008 and 2009 (and divide weights by two) so that the uninsured sample reflects an average of the two relevant calendar years. We restrict to Massachusetts residents age 19-64, since seniors are in Medicare and low-income children (up to 300% of poverty) are in Medicaid in Massachusetts. We restrict to people with household income $\leq 300\%$ of the federal poverty level. We define households as "health insurance units," a variable included on the IPUMS ACS that is intended to approximate the household definition used by public insurance programs. We exclude non-citizens because most are ineligible, and the rare exception (long-term green card holders) cannot be measured. We also exclude the uninsured who are eligible for Medicaid (rather than CommCare) – parents up to 133% poverty and disabled individuals (proxied by receiving SSI income). We correct for the fact that the ACS is a sample by weighting ACS observations by their "person weight" – the Census-defined factor for scaling up to a population estimate.

Although focusing on new enrollees simplifies the demand model, it creates an additional complication in combining the ACS and CommCare data. While we can differentiate new and existing consumers in the CommCare data, the ACS only lets us observe the total *stock* of uninsured in each year. To convert this into a comparable flow of "new uninsured," we adjust the ACS data weights so that the uninsured share in our final demand sample matches the overall eligible population uninsurance rate. Specifically, we first estimate the uninsurance rate in each year using *all individuals* in both datasets – the population estimate of the number of CommCare-eligible uninsured from the ACS and the number of covered individuals (in member-years) in the CommCare data (including both new and current enrollees). We then rescale the ACS weights in our demand estimation sample – which contains all ACS uninsured observations but only new enrollees for the CommCare data – so that the uninsured rate calculated in this sample matches the population uninsured rate.

B.2 Natural Experiments

This section provides more background on the natural experiments we use and details of the estimation as well as some robustness checks.

Mandate Penalty Introduction

As described in the text, the mandate penalty went in to effect in December 2007. We estimate excess new enrollments in December 2007-March 2008 relative to the trend in nearby months, using enrollment trends for people earning less than poverty as a control group. We estimate the effect through March 2008 for two reasons. First, the application process for the market takes some time, so people who decided to sign up in January may not have enrolled until March. Second, the mandate rules exempted from penalties individuals with three or fewer months of uninsurance during the year, meaning that individuals who enrolled in March avoided any penalties for 2008. However most of the effect is in December and January, so focusing on those months does not substantially affect our estimates.

We collapse the data to the income group-month level and calculate the new enrollees

in the cheapest plan for each group and month, normalized by that plan's total enrollment for the income group in June 2008 (a proxy for the steady-state size of each plan). We use data only up to June 2011 because of significant changes in the prices and availability of the cheapest plans that took effect in July 2011. We estimate the difference-in-difference equation shown in in Section III.A:

$$NewEnroll_{y,t} = (\alpha_{0,t} + \alpha_{1,t} \cdot Treat_y) \cdot MandIntro_t + \xi_y + \delta_y \cdot X_t + \zeta_{y,m(t)} \cdot DM_t + \varepsilon_{y,t}.$$

The difference-in-difference coefficients of interest are the $\alpha_{1,t}$'s. In our main specifications, we use two income groups (y): 150-300% as the treatment group and below 100% as the control group. We also break down estimates for the treatment group separately by 50% of poverty group. People earning 100-150% of poverty are omitted from the control group because a large auto-enrollment took place for this group in December 2007, creating a huge spike in new enrollment. But the spike occurred only in December and was completely gone by January, unlike the pattern for the 150-300% poverty groups. This auto-enrollment did not apply to individuals above 150% of poverty (Commonwealth Care, 2008) so it cannot explain the patterns shown in Figure 1.

Table 1 presents the regression results. Column (1) starts with a baseline single-difference specification (i.e., without the control group) that estimates the effect based only on enrollment for the 150-300% poverty group in December 2007-March 2008 relative to the surrounding months. Column (2) then adds the $< 100\%$ poverty group as a control group, to form difference-in-difference estimates. Finally, Column (3) adds dummies for December-March in all years, forming the triple difference specification (shown in the equation above) that nets out general trends for those months in other years. The final three columns take the final, triple-difference specification and breaks down the analysis by limiting the treatment group to narrower income groups (but keeping the control group, $< 100\%$ poverty, unchanged).

Despite the relatively small number of group-month observations, all the relevant coefficients are statistically significant – consistent with the dramatic spike shown in Figure 1. In our preferred triple difference estimates in column (3), the mandate penalty increases enrollment in the cheapest plan by 22.2% of its steady state size. When we break these results down by income group, the coefficients are slightly larger for higher income groups – about 25% instead of 21% – who faced higher mandate penalties. The mandate penalties as of January 2008 were $17.50 for the 150-200% poverty group, $35 for the 200-250% poverty group, and $52.50 for the 250-300% poverty group. Given these penalties, we can use our estimates to calculate semi-elasticities. The estimates imply that each $1 increase in the mandate penalty raised demand by 1.17% for the 150-200% of poverty group, 0.74% for the 200-250% poverty group, and 0.47% for the 250-300% of poverty group, with a (group-size) weighted average of 0.95%.

We interpret the increases in enrollment as being the result of the permanent $17.50– $52.50 monthly mandate penalty that went into effect in January 2008. However the 2007 uninsurance penalty – forfeiting the state tax personal exemption, with a value of $219 – was assessed based on coverage status in December 2007, making that month's effective penalty much larger (though the total *annual* penalty for 2008 was actually larger than in 2007 for all but the 150-200% poverty group). Technically, individuals who applied for CommCare in 2007 and were enrolled on January 1, 2008, did not owe the penalty for 2007. But since

Table 1: Introduction of the Mandate Penalty.

Effect on New Enrollees in Cheapest Plan / June 2008 Enrollment

	All Treatment Groups			Income Group (% of Poverty)		
	(1)	(2)	(3)	150-200%	200-250%	250-300%
Dec '07	0.112	0.110	0.102	0.101	0.112	0.091
	(0.003)	(0.006)	(0.007)	(0.007)	(0.007)	(0.008)
Jan '08	0.073	0.067	0.067	0.059	0.084	0.086
	(0.004)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
Feb '08	0.043	0.033	0.032	0.026	0.042	0.049
	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)
Mar '08	0.025	0.027	0.021	0.020	0.021	0.022
Total	**0.253**	**0.237**	**0.222**	**0.205**	**0.259**	**0.248**
	(0.018)	(0.023)	(0.024)	(0.026)	(0.023)	(0.023)
Controls:						
<100% Poverty		X	X	X	X	X
Dec-Mar			X	X	X	X
Observations	51	102	102	102	102	102
R-squared	0.967	0.921	0.923	0.925	0.920	0.919

Robust s.e. in parentheses

Note: This table performs the difference-in-difference regressions analogous to the graphs in Figure 1. The dependent variable is the number of new CommCare enrollees who choose the cheapest plan in each month in an income group, scaled by total group enrollment in that plan in June 2008. There is one observation per income group and month (from April 2007 to June 2011). All specifications include CommCare-year dummy variables and fifth-order time polynomials, separately for the treatment and control group. (The CommCare-year starts in July, so these dummies will not conflict with the treatment months of December to March.) Columns (2) adds the < 100% poverty group as a control. Column (3) also includes dummy variables for the calendar months of December-March in all years, separately for the treatment and control group, to perform the triple-difference. Columns (4)-(6) do the same regression as Column (3) separately by income group. See the note to Figure 1 for the definition of new enrollees and the cheapest plan.

10

December 31, 2007, was the main advertised date for assessing the 2007 penalty, we want to make sure that the larger effective penalty for that month is not driving the results.

To test this, we note that if consumers were only buying insurance because of the larger December penalty, we would expect many of them to leave the market soon after the monthly penalty dropped to the lower level in January 2008. Figure 1a and 1b plot the probability of exiting the market within 1 month and 6 months of initial enrollment for each entering cohort of enrollees. The graph shows that exit probabilities are no higher for people entering CommCare in December 2007 than for nearby months. (Note that the large spike in one-month exits for the March 2008 cohort is due to an unrelated income verification program.)[4] This analysis suggests that consumers were not enrolling for just December to avoid the larger penalty and leaving soon afterward.

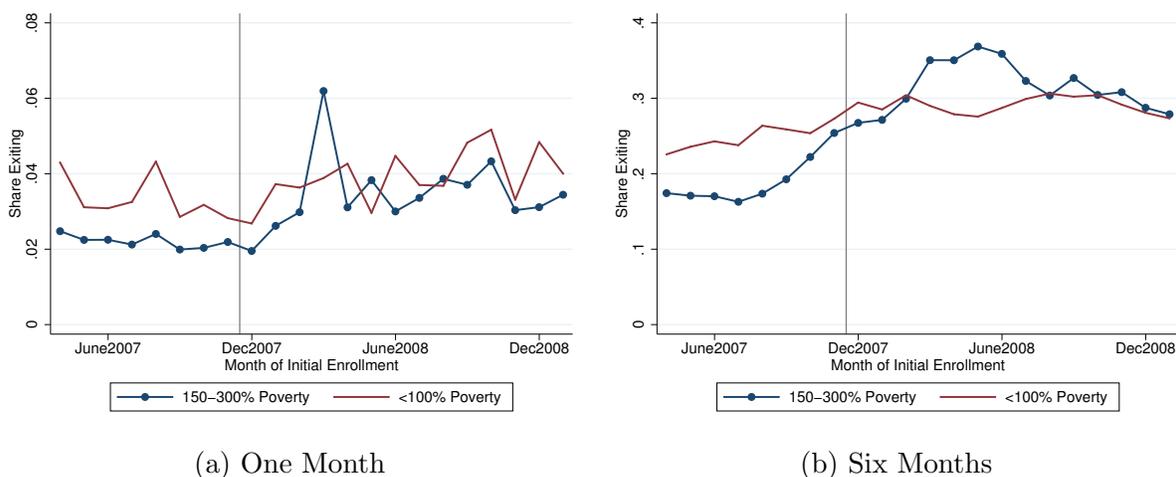(a) One Month   (b) Six Months

Figure 1: Share of New Enrollees Exiting Within the Specified Number of Months.

Note: These graphs show the rate of exiting CommCare coverage within (a) one month and (b) six months of initial enrollment among people newly enrolling in CommCare in a given month. The spike among new enrollees in March 2008 reflects the start of an income-verification program for the 150-300% poverty group in April 2008. See the note to Figure 1 for the definition of new enrollees and the cheapest plan.

Premium Decrease Experiment

Our second natural experiment addresses a potential concern with our first method: that the *introduction* of a mandate penalty may have a larger effect (per dollar of penalty) than a *marginal increase* in penalties. Some individuals may obtain coverage to avoid the stigma of paying a penalty, but this stigma might not change when mandate penalties increase. An argument against the stigma explanation is that the legal mandate to obtain insurance had been in place since July 2007 and also applied to the control group (but without financial enforcement). However to the extent there is a stigma specifically from paying a fine for non-coverage, this concern is valid.

---

[4]The income verification program took effect in April 2008 for individuals above 150% of poverty. The program uncovered a large number of ineligible people, who were dis-enrolled in April 2008 and subsequent months. This also explains the upward trend in exits within 6 months leading up to April 2008.

Our second experiment is a decrease in the enrollee premium of all plans – via a reduction in the "affordable amount" that determines the post-subsidy premium of the cheapest plan – that occurred in July 2007. Although prices and contracts were fixed from the start of CommCare (in November 2006) until June 2008, the government decided to increase subsidies for certain groups in July 2007 to make insurance more affordable. For the 150-200% of poverty group, the affordable amount fell from $40 to $35 – which meant that the monthly premium of all plans fell by $5. For consumers between 100-150% of poverty, the affordable amount was $18 for the first half of 2007 and premiums were $18-$74.22. In July 2007, CommCare eliminated premiums for this group, so all plans became free. We can think of this as the combination of two effects: (1) The affordable amount was lowered from $18 to zero, and (2) the premium of all plans besides the cheapest one were differentially lowered to equal the cheapest premium (now $0). The second change should unambiguously lower enrollment in what was the cheapest plan, since the relative price of all other plans falls. So the aggregate effect of these changes is a lower bound on the effect of just lowering the affordable amount.

As a control group, we use the 200-300% poverty group, whose affordable amounts were essentially unchanged in July 2007.[5] We exclude the below 100% poverty group from our controls because of its somewhat different enrollment history and trends. Whereas the groups above poverty only started joining CommCare in February 2007, the below 100% poverty group became eligible in November 2006 and had a large influx in early 2007 due to an auto-enrollment. Figure 2 shows enrollment in the cheapest plan as a share of June 2008 enrollment for the 100-150% poverty treatment group and the 200-300% poverty control group as well as other years for the treatment group. There is a jump in treatment group enrollment in July 2007. (The jump in the control group in January 2008 is due to the introduction of the mandate penalty.)

Table 2 presents the regression results. In Columns (1) and (4) we look at the single difference for the treatment group relative to trend (i.e., without a control group). Column (2) and (5) then add the 200-300% poverty control group to form the difference-in-difference estimates. Columns (3) and (6) do a triple-difference, further netting out changes in July-October of other years. In this triple difference specification, for enrollees 100-150% of poverty, we find a 17.4% increase in the cheapest plan's demand. Dividing by the $18 reduction in its premium implies a semi-elasticity of 0.97%. For the 150-200% poverty group, we find a 6.3% increase in demand. Dividing by its $5 premium reduction yields a semi-elasticity of 1.26%. We note that this semi-elasticity is very close to the 1.17% semi-elasticity for the 150-200% poverty group found in the mandate penalty introduction experiment.

Own-price Semi Elasticity

As described in Section III.B, we use within plan variation in consumer premiums generated by the subsidy rules to estimate the semi-elasticity of demand with respect to own-price. Figure 3 shows the same changes as Figure 2 in the text, but for plans that increase their

---

[5]The affordable amount for 200-250% poverty was unchanged and that for 250-300% poverty was lowered by just $1 from $106 to $105; to the extent this slightly increased enrollment for the control group, it would bias our estimates downward.

Table 2: Decrease in the Affordable Amount.

Effect on New Enrollees in Cheapest Plan / June 2008 Enrollment

	100-150% Poverty			150-200% Poverty		
	(1)	(2)	(3)	(4)	(5)	(6)
July '07	0.065	0.061	0.065	0.014	0.017	0.019
	(0.005)	(0.008)	(0.009)	(0.005)	(0.007)	(0.008)
Aug '07	0.053	0.032	0.036	0.029	0.016	0.022
	(0.005)	(0.008)	(0.009)	(0.005)	(0.007)	(0.008)
Sep '07	0.022	0.011	0.017	0.013	0.011	0.016
	(0.005)	(0.008)	(0.008)	(0.005)	(0.006)	(0.007)
Oct '07	0.063	0.054	0.057	0.008	0.009	0.010
	(0.005)	(0.008)	(0.008)	(0.005)	(0.006)	(0.006)
Total	**0.202**	**0.157**	**0.174**	**0.063**	**0.052**	**0.063**
	(0.021)	(0.031)	(0.032)	(0.019)	(0.025)	(0.027)
Controls:						
200-300% Poverty		X	X		X	X
Jul-Oct '08-'11			X			X
Observations	52	104	104	52	104	104
R-Squared	0.986	0.980	0.980	0.958	0.955	0.956

Robust s.e. in parentheses

Note: This table reports difference-in-difference regressions of the effect of the decrease in the affordable amount. The dependent variable is the number of new CommCare enrollees who chose the cheapest plan, scaled by total group enrollment in that plan in June 2008. There is one observation per income group and month (from March 2007 to June 2011). All specifications include CommCare-year dummy variables and fifth-order time polynomials, separately for the treatment and control group, to control for underlying enrollment trends. (The first CommCare year ended in June 2008, so there is no conflict between the CommCare-year dummies and the treatment months of July to October 2007.) Columns (2) and (5) add the 200-300% poverty control group. Columns (3) and (6) add controls for July-October of other years, separately by treatment and control group. Where applicable, specifications also include dummy variables to control for two unrelated enrollment changes: (a) for 100-150% poverty in December 2007, when there was a large auto-enrollment spike, and (b) for 200-300% poverty in each month from December 2007 to March 2008, when there was a spike due to the mandate penalty introduction. See the note to Figure 1 for the definition of new enrollees and the cheapest plan.
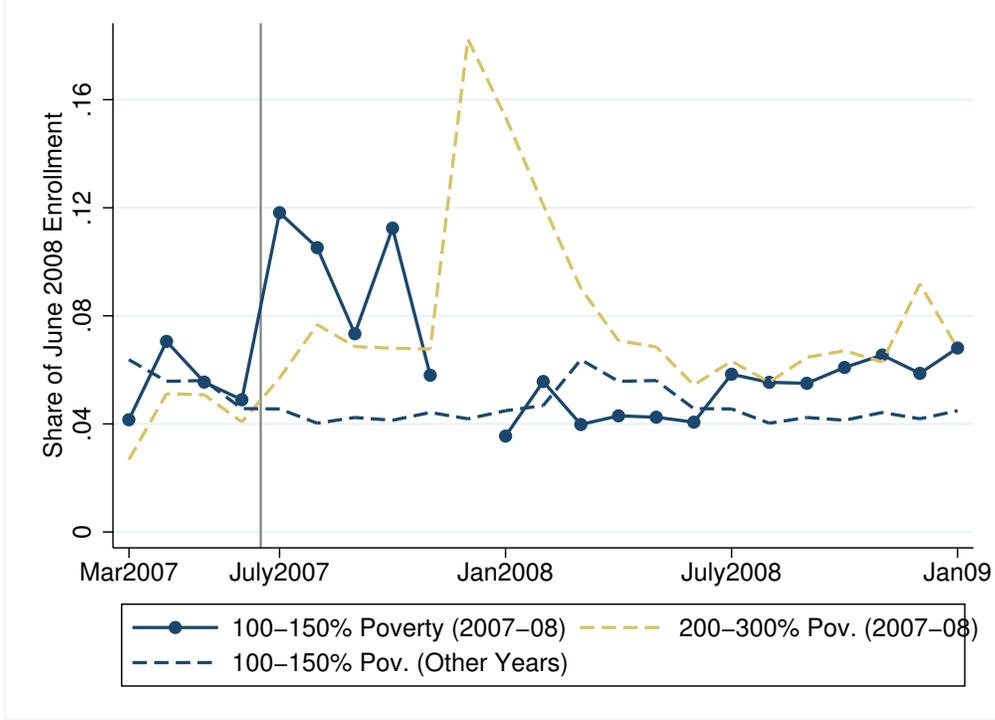
Figure 2: Decrease in the Affordable Amount Experiment

Note: Analagously to Figure 1, this graph shows monthly new enrollees (both first-time consumers and those re-enrolling after a break in coverage) into CommCare's cheapest plan as a share of total June 2008 enrollment, so units can be interpreted as fractional changes in enrollment for each group. The vertical line is drawn just before the decrease in the affordable amount, which affected the "100-150% Poverty" income group. The "100-150% Poverty (Other Years)" combines all years in our data *except* July 2007–June 2008. The spike in the control group ("200-300% Pov") in December 2007 is due to the introduction of the mandate penalty (our other natural experiment), and we dummy out Dec 2007 to March 2008 for the control group so that this does not affect the regression estimates. The excluded observation for the treatment group is due to a large auto-enrollment that happened for that group that month.

prices. Table 3 reports the coefficient on premium (η) from the regression in Equation (6):

$$\ln\left(NewEnroll_{j,y,r,t}\right) = \eta \cdot P^{cons}_{j,y,r,t,} + \xi_{j,r,t} + \xi_{j,r,y} + \epsilon_{j,y,r,t}.$$

The first column has no controls. Since price and quality tend to be positively correlated, it is not surprising that adding plan-region-income and plan-region-year controls (second column) increases the coefficient by about 50%. Since the regions in Massachusetts are different sizes, the last column weights by the region's average monthly new enrollment, generating our semi-elasticity estimate of 2.16% used in the calculations.
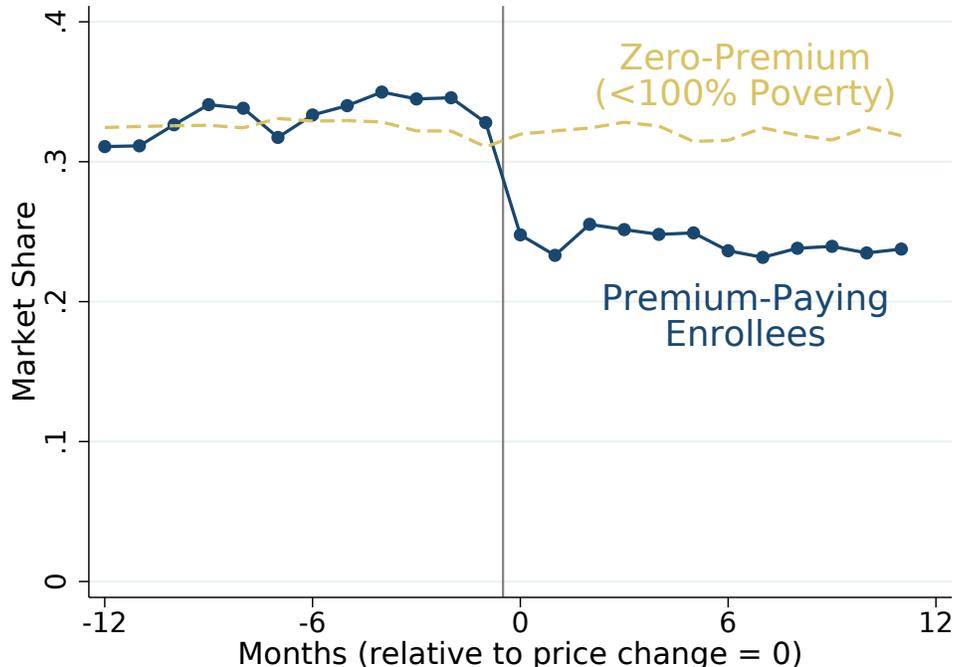
Figure 3: Market share around price increases

Note: This graph illustrates the source of identification for the estimation of the own-price semi-elasticity of demand. The graph shows average monthly plan market shares among new enrollees for plans that increased their prices at time 0. The identification comes from comparing demand changes for above-poverty price-paying (new) enrollees (for whom premium changes at time 0) versus below-poverty zero-price enrollees (for whom premiums are always $0). The sample is limited to fiscal years 2008-2011, the years we use for demand estimation.

C   Structural Model Details and Additional Results

C.1   Moments and Method Details

Demand Model

We estimate the model by simulated method of moments, incorporating micro moments with an approach similar to Berry, Levinsohn and Pakes (2004). We draw a ν_i from the N(0,1) distribution for each individual i (with their associated Z_i) in the demand estimation sample (using data from both the CommCare and ACS data; see Table 1). Given logit errors, the plan choice probabilities are

$$P(j_i^* = j | Z_{it}, \nu_i, \theta) = \frac{\exp(\hat{u}_{ij})}{\sum_{k=0}^{J} exp(\hat{u}_{ik})}.$$

where θ refers to the parameters to be estimated and $\hat{u}_{ij} = u_{ij} - \epsilon_{ij}$ is the systematic part of utility.

For each ξ (the CommCare plan utility coefficients) we have a moment for the corre-

Table 3: Differential Premium Changes.

Effect on Price on Log New Enrollment

	(1)	(2)	(3)
Premium	-0.0129	-0.0193	-0.0216
	(.00077)	(.00168)	(.00175)
Plan X region X income dummies		X	X
Plan X region X year dummies		X	X
Weight: Region avg enrollment			X
Observations	4713	4713	4713
R-Squared	0.238	0.891	0.882

Robust s.e. in parentheses

Note: This table reports difference-in-difference regressions of the effect of a plan's premium on the number of new enrollees for 2008 through 2011. Column (1) has no controls. Column (2) adds plan-region-year dummies and plan-region-income dummies. Enrollees below 100% of the federal poverty line pay zero premium for all plan, so they control for changes in plan quality over time. Other enrollees pay higher premiums for more expensive plans. Column (3) weights by the region's average new enrollees each month. Standard errors are clustered at the region-income-year level.

sponding plan and group, g (either region-year or region-income group) that matches the expected share of consumers in that group who chose that plan given θ, to the observed share, s_j^{Obs}. If n_g is the number of individuals in group g, we have

$$F_{j,g}^1(\theta) = s_{j,g}^{Obs} - \frac{1}{n_g} \sum_{i \in g} Pr(j_i^* = j | \theta, Z_i, \nu_i.)$$

For each β (in the utility of uninsurance) there is also a corresponding group, h – income, demographic, region, or year. Analogous to the moments in F^1, we match the observed share of uninsured to that predicted for the model. We do it separately by age-gender groups, income groups, regions, and year, but because the uninsured data are based on relatively small samples in the ACS, we do not interact these categories. The corresponding moments are

$$G_{0,h}^1(\theta) = s_{0,h}^{Obs} - \frac{1}{n_h} \sum_{i \in h} Pr(j_i^* = j | \theta, Z_i, \nu_i).$$

To identify the different price-sensitivity parameters (α's), we match the covariance of plan premium and individual attributes, following Berry, Levinsohn and Pakes (2004). For each income or demographic group, h, we use

$$G_h^2(\theta) = \frac{1}{n} \sum_j \sum_i P_{ij}^{Cons} Z_i \left(1\{j_i^{Obs} = j\} - Pr(j_i^* = j | \theta, Z_i, \nu_i) \right)$$

16

with $P_{i,0}^{Cons} = M_i$.

The final moment helps identify the variance of the random coefficients by matching the estimated insurance demand response from the natural experiments discussed in Section II.A. If there is substantial heterogeneity in the value of insurance, the uninsured will tend to be people with very low idiosyncratic values of insurance; since they are not close to the margin of buying coverage, an increase in the mandate penalty will not increase demand for insurance very much. Thus, higher values of σ are likely to generate less demand response to the mandate penalty, and vice versa. We match the simulated change demand for the cheapest plan to the observed 22.2% change in demand for the mandate penalty introduction experiment:

$$G^3(\theta) = \sum_i \left((1 + 22.2\%) Pr(j_i^* = j_{min} | Z_i, \nu_i, \theta, M_i^{Pre}) - Pr(j_i^* = j_{min} | Z_i, \nu_i, \theta, M_i^{Post}) \right),$$

where the first probability in the equation is based on the pre-period mandate penalty (M_i^{Pre}) – which is zero – and the second probability is based on the post-period mandate penalty (M_i^{Post}) that applied from January 2008 on. We do not use the premium decrease experiment estimates because it occurred before the start of our demand estimation period (January 2008). However, because the semi-elasticity estimated from it is so similar, including it as a moment would be unlikely to affect our results.

The model is just-identified, and our parameter search exactly matches the empirical moments.

Cost Model

The cost model we use is called a GLM with a Poisson error and a log link function; it is equivalent to a Poisson regression, although allowing for a continuous outcome variable. Mathematically, the model is estimated by MLE as if the data were generated by a Poisson process with mean equal to the expression in Equation (7). However, in practice, this is simply a technique to estimate the parameters in Equation (7), which describe the conditional mean of costs. These conditional means are what we use for costs in our simulations; we do not use the Poisson distribution. This method has the advantage over a log-linear regression of allowing for $0 observations, which occur regularly in our data.

A concern with a basic maximum likelihood estimation of the cost model is that estimates of $\psi_{j,r,t}$ will be biased by selection on unobserved sickness. This is particularly relevant because X_i includes a relatively coarse set of observables. To partially address this issue, we estimate the $\psi_{j,r,t}$ parameters in a separate version of Equation (7) with individual fixed effects and a sample limited to new and re-enrollees. These estimates are identified based only on *within-person* cost variation when an individual leaves the market and later re-enrolls in a different plan (e.g., because plan prices have changed). This method would fully eliminate selection if individuals' unobserved risk factors are stable over time or uncorrelated with plan changes. However, it would not address any selection on risk changes – e.g., if individuals who get sicker between enrollment spells systematically select certain plans.[6]

---

[6]Our sense is that controlling for individual fixed effects addresses most of the selection problem, especially since our panel is over a short period (about four years).

Lastly, we adjust observed costs by removing the estimated plan component and estimate the coefficients on individual characteristics from cross-person variation with all enrollees. We use the resulting predicted values of $E(c_{ijt})$ as our estimates of costs for each enrollee-plan possibility.[7]

C.2   Parameter Estimates

The demand model coefficients are presented in Tables 4, 5, and 6. Table 4 shows that on average, consumers prefer Fallon to BMC to Network Health to NHP to CeltiCare, with the biggest difference being between CeltiCare and the other plans.

Table 4: Demand Estimates: Average Plan Coefficients

	Coef	S.E.
CeltiCare	-0.857	0.0564
NHP	-0.067	0.0147
Network	0.126	0.0108
BMC	0.169	0.0312
Fallon	0.209	0.0563

Note: The table reports the average exchange plan dummies (with the average plan normalized to 0); these are weighted averages of the full set of plan-region-year and plan-region-income group dummies in the model.

Table 5 shows the baseline price coefficients by income and demographic group. The income coefficients (which apply to the omitted demographic group of males age 40-44), range from -0.046 for those just above the poverty line to -0.023 for those making 250%-300% FPL. Females and older consumers are less price sensitive; on average, the premium coefficient increases by +.0028 for women and +.0019 for each 5 years of age. The oldest group is less than half as price-sensitive as the youngest. (Every income group has a premium coefficient < -0.02, and no demographic group's coefficient is $> .012$, so every group has a negative total premium coefficient.)

Table 6 shows the utility of uninsurance (β), which is relative to the utility of the average plan (which we normalized to 0), by income and demographic group. Females and older people dislike uninsurance more. The average relative utility of uninsurance for above-poverty consumer (not shown in table) is positive (+0.261), which is necessary to account for the low insurance take-up rates in our data despite very low (and often zero) subsidized premiums. We also note that the unobservable variation (σ) in the value of uninsurance is substantial; it accounts for a standard deviation of 0.92 across individuals, relative to the 0.67 standard deviation driven by observables.

Table 7 shows the average cost differences across income groups. Table 8 shows the cost differences across age and gender groups.

---

[7]In addition to insurer medical costs captured by this model, insurers incur administrative costs for functions like claims processing. Using plan financial reports, we estimate variable administrative costs of approximately $30 per member-month. We add this to our cost function, and following CommCare rules, insurers are also paid an equal-size ($30) administrative fee (not subject to risk adjustment). Thus, administrative costs cancel out in the profit function and do not impact our results.

Table 5: Demand Estimates: Interaction terms for Premium Coefficients

Baseline by income group (For omitted group: males 40-44)			Interacted with demographic group (To be added to baseline value)				
				Male		Female	
	Coef	S.E.		Coeff	S.E.	Coeff	S.E.
100-150% Pov	-0.046	0.0054					
150-200% Pov	-0.029	0.0061	Age 19-24	-0.0086	0.0049	0.0011	0.0026
200-250% Pov	-0.027	0.0072	Age 25-29	-0.0097	0.0047	-0.0003	0.0033
250-300% Pov	-0.023	0.0010	Age 30-34	-0.0012	0.0028	-0.0048	0.0042
			Age 35-39	-0.0016	0.0033	-0.0016	0.0039
			Age 40-44			0.0013	0.0035
			Age 45-49	0.0077	0.0037	0.0006	0.0037
			Age 50-54	0.0050	0.0030	0.0013	0.0034
			Age 55-59	0.0122	0.0037	0.0084	0.0034
			Age 60-64	0.0111	0.0038	0.0164	0.0037

Note: This table shows the the premium coefficient by income (left) and for each gender x age bin (right). The full premium coefficient for any sub-group is the sum of the coefficients for the corresponding income group and demographic group. Every income group has a premium coefficient < -0.02, and no demographic group's coefficient is $> .012$, so every group has a negative total premium coefficient.

Table 6: Demand Estimates: Relative Utility of Uninsurance

Baseline by income group (For omitted group: males 40-44)			Interacted with demographic group (To be added to baseline value)				
				Male		Female	
	Coef	S.E.		Coeff	S.E.	Coeff	S.E.
<100% Pov	0.952	0.142					
100-150 % Pov	0.667	0.174	Ages 19-24	0.236	0.109	-0.596	0.201
150-200 % Pov	0.150	0.319	Ages 25-29	0.095	0.117	-0.792	0.240
200-250 % Pov	0.042	0.526	Ages 30-34	0.066	0.128	-0.889	0.259
250-300 % Pov	0.517	0.825	Ages 35-39	-0.067	0.162	-0.844	0.251
			Ages 40-44			-1.073	0.324
			Ages 45-49	-0.043	0.135	-1.097	0.329
			Ages 50-54	-0.229	0.156	-0.874	0.251
			Ages 55-59	-0.299	0.161	-0.778	0.256
			Ages 60-64	-0.590	0.253	-0.785	0.298

Note: This table shows the the relative utility of uninsurance by income (left) and for each gender x age bin (right). The total average utility of uninsurance (relative to the average CommCare plan) for a group is the sum of the coefficients for the corresponding income group and demographic group. Some groups have a positive utility of uninsurance (negative value of insurance), which is necessary to explain the fact that there are many people for whom insurance is free who nonetheless choose not to purchase insurance.

Table 7: Cost by income group

| | Income Group Percent of Federal Poverty Line | | | |
	100-150	150-200	200-250	250-300
Relative to < Poverty	-23.4%	-18.2%	-25.2%	-29.5%

Note: This table reports the costs for each income group relative to the below poverty group. The reported percentage for group g is $100\cdot(\exp(\mu)-1)$, when the μ_g for the 'Less than Poverty' group is set to zero.

Table 8: Costs by Age and Gender

| | Male | | | Female | | |
	Poisson Coefficient	S.E.	Percent	Poisson Coefficient	S.E.	Percent
Ages 19-24				0.312	0.00031	36.6%
Ages 25-29	0.336	0.00032	39.9%	0.488	0.00028	62.9%
Ages 30-34	0.424	0.00035	52.8%	0.558	0.00031	74.7%
Ages 35-39	0.526	0.00035	69.2%	0.609	0.00036	83.9%
Ages 40-44	0.632	0.00033	88.1%	0.752	0.00036	112.1%
Ages 45-49	0.807	0.00031	124.1%	0.850	0.00032	134.0%
Ages 50-54	0.943	0.00030	156.8%	0.886	0.00030	142.5%
Ages 55-59	1.065	0.00030	190.1%	0.907	0.00029	147.7%
Ages 60-64	1.193	0.00031	229.7%	1.073	0.00030	192.4%

All coefficients are significant at the $p < .01$ level.

Note: This table shows cost coefficients by age and gender. Males 19-24 are the omitted group. For each gender, the first two columns are the Poisson coefficient and its standard error. The third coefficient converts this to a percent difference, $100\cdot(\exp(\mu)-1)$.

C.3   Simulation Method Details

C.3.1   Risk Adjustment Method

To mitigate adverse selection, both Massachusetts and the ACA risk-adjust payments to insurers. We incorporate this by estimating a risk score for each individual in each year, ϕ_{it}, that indicates how costly they are expected to be relative to the average enrollee. Following Massachusetts' rules, a plan with price P_{jt} receives $\phi_{it}P_{jt}$ for enrolling consumer i in year t. We assume that ϕ_{it} perfectly captures the individual (non-plan) portion of costs: $\phi_{it} = \exp\left(\hat{\mu}X_{it} + \psi_{0t}\right)/\overline{c}_t$, where $\overline{c}_t$ is the average of $\exp\left(\hat{\mu}X_{it} + \psi_{0t}\right)$ across individuals who buy insurance in the data. Thus, risk adjustment captures individuals' expected costs in a proportional way: net revenue on consumer i ($= \phi_{it}P_{jt} - E(c_{ijt})$) simplifies to $\phi_{it}\left(P_{jt} - \overline{c}_t \exp(\psi_{j,r,t})\right)$.

C.3.2   Equilibrium

We look for a static, full information, Nash equilibrium, where each insurer sets its price to maximize profits:

$$\pi_{jt} = \sum_i \left(\phi_{it}P_{jt} - E[c_{ijt}]\right) \cdot Q_{ijt}\left(P^{Cons}\left(P_t\right)\right) \tag{C.1}$$

where $P^{Cons}\left(\cdot\right)$ is the subsidy function mapping prices into consumer premiums. This equilibrium is defined by the first-order conditions (FOCs) $d\pi_{jt}/dP_{jt} = 0$ for all j, given all other plans' prices. For each year and subsidy policy, we simulate the equilibrium numerically by searching for the price vector P that satisfies these equilibrium conditions for all insurers. We do this separately for the set of insurers present in the market in 2009 and 2011 and for counterfactual markets with only two insurers.

Note that the insurers' FOCs depend on $\frac{dQ_{jt}}{dP_{jt}}$, the total derivative that incorporates an effect of changing P_{jt} on the subsidy if plan j is pivotal (cheapest) under price-linked subsidies (as shown in Equation (1)). This introduces a discontinuity in the FOC of the cheapest plan at the price of the second-cheapest plan: below it, they are subsidy-pivotal (so $\frac{dQ_{jt}}{dP_{jt}} = \frac{\partial Q_{jt}}{\partial P_{jt}} + \frac{\partial Q_{jt}}{\partial M}$) while above it, they are not (so $\frac{dQ_{jt}}{dP_{jt}} = \frac{\partial Q_{jt}}{\partial P_{jt}}$). This means that if the incentive distortion causes the cheapest plan to raise its price up to the level of the second cheapest plan, there will likely be multiple equilibria – a range of prices at which the two plans tie for cheapest in equilibrium. In our simulations this occurs in 2009, so we show results for both the minimum and maximum prices consistent with this range of equilibria.

C.3.3   Regulator's Objective (Welfare Function) for Simulations

Our main welfare analysis focuses on "public surplus," introduced in Section I.B – consumer surplus, plus the avoided externality of uninsurance, minus government costs. See Appendix A.4 for the formula for public surplus. We also consider a version that adds insurer profits.

For each simulation, we use the simulated equilibrium prices along with the estimated demand parameters to calculate the fixed component of utility for each consumer for each plan, $\hat{u}_{ij} \equiv u_{ij} - \epsilon_{ij}$. With logit demand, these allow us to calculate each individual's

21

expected consumer surplus $CS_i = \frac{1}{-\alpha(Z_i)} \cdot \log \left(\sum_{j=0}^{J} \exp(\hat{u}_{ij}) \right)$ and choice probabilities $\left(\widehat{Pr}_{ij} = \exp(\hat{u}_{ij}) / \sum_{j=0}^{J} \exp(\hat{u}_{ij}) \right)$. The government's expected net expenditure on each individual is $S(1 - \widehat{Pr}_{i0}) - M \cdot \widehat{Pr}_{i0}$.

Uninsured people can generate a negative externality through the uncompensated care that they receive (whose cost is born by the providers or government); if society or the regulator has a paternalistic desire for people to have health insurance, then that is another negative effect of people being uninsured, independent of the medical costs they incur. We assume that the regulator sets the affordable amounts for price-linked subsidies optimally given the total externality. This allows us to back out the level of the paternalistic externality.[8] This calibration is conceptually important. Without the paternalistic externality, subsidies are "too high" from a welfare-maximizing perspective so there are "too many" people buying insurance. This makes fixed subsidies seem bad because they lead to more people buying insurance, whereas we want to compare fixed and price-linked subsidies assuming they are set at the right level, given the externality.

C.4 Simulation Results

C.4.1 Comparison of reduced form and structural distortion estimates

Our paper includes two main estimates of the distortion from price-linked subsidies: a "reduced form" estimate of $36.35 based on the first-order approximation formula (reported in Section III.C) and a "structural" estimate of $23.83 from our simulation model for 2011 (reported in Section IV.D). In this appendix, we seek to separate out how much of the difference is due to (1) different estimated elasticities, (2) adverse selection and risk adjustment, (3) pass-through of costs into prices not equaling 1 (which is implicitly assumed by the first-order approximation), and (4) strategic responses from the other firms.

Table 9 breaks down the role of each of these factors. Line (a) shows our reduced form estimate of $36.35, while the final line (e) is the headline structural estimate of $23.83 for 2011. Bridging these two estimates are the following factors:

- **Different time period (line (b)):** The reduced form estimate uses elasticities from 2008 (when the natural experiment occurred), while the headline structural estimate is from 2011. Line (b) takes demand elasticities for 2011 simulated from our structural model and plugs these into the reduced form distortion formula. The result is a distortion of $30.02 – bridging about half of the original gap. From this we conclude that the differing time periods is a major source of the gap.

- **Incorporating adverse selection (line (c)):** Our main distortion formula ignores adverse selection, but the structural model allows for it. To incorporate adverse selection, we use an augmented distortion formula derived in Appendix A.2 (Equation

---

[8]For these calculations, we assume that uncompensated care costs are equal to 80% of expected cost of care with insurance and that the paternalistic component of the externality does not vary with consumers' expected costs. See Section V.A for further discussion of uncompensated care costs and Appendix D for robustness checks with the externality equal to 60% and 100% of expected cost.

(A.1)). Simulating the necessary additional statistics using our structural model and plugging these into the formula, we get a distortion of $26.71 (bridging another fourth of the original gap).

- **Moving to a structural simulation (no strategic interactions) (line (d)):** This line implements a structural simulation of the distortion for CeltiCare in 2011 that shuts down strategic pricing responses by other firms. The difference is accounted for by the fact that the reduced form estimate is a first-order approximation (which, e.g., implicitly assumes constant semi-elasticities of demand and a pass-through rate of 1). The estimate in line (d) is $24.34, closing another 17% of the original gap.

- **Adding strategic interactions:** Line (e) shows our final structural simulation that allows for strategic pricing responses by competitors. The difference between (d) and (e), however, is tiny (just $0.51), indicating that strategic effects explain very little of the original gap.

Table 9: Comparison of reduced form and structural estimates

Calculation method		Estimate ($)
Reduced-form formula		
(a)	Baseline	36.35
(b)	With structural model elasticities	30.02
(c)	+ Adverse selection	26.71
Structural model		
(d)	Without strategic interactions	24.34
(e)	Full model	23.83

C.4.2 Medical Loss Ratios

Regulatory limits on the minimum fraction of premiums that must be used to cover medical care could limit the distortionary impact of price-linked subsidies. The ACA, sets a minimum medical loss ratio (MLR) of 80%. Table 10 shows estimated costs as a fraction of revenue for each firm for each policy simulation. The only MLR that is below the 80% threshold is BMC's under the maximum-price equilibrium for price-linked subsidies in 2009. This suggests MLR rules may have had some bite for avoiding the more expensive equilibrium in 2009, but would not have limited the distortion 2011.

When there are only two insurers, profits are higher and MLR rules become more binding. Table 11 shows the estimated MLRs if there were only two insurers in the market in 2011. CeltiCare's MLR is below 80% for price-linked subsidies whenever it has only one competitor and for fixed subsidies when it competes against NHP (the highest cost insurer). Network Health and BMC also have MLRs below 80% when competing against NHP. MLR rules are more relevant for extremely uncompetitive markets.

Table 10: Medical Loss Ratios

		BMC	Celticare	Fallon	NHP	Network Health
2009	Min Price-linked	81.5%		91.2%	91.4%	89.9%
	Max Price-linked	77.2%		92.1%	92.1%	85.3%
	Fixed	82.3%		91.2%	91.3%	90.1%
2011	Price-Linked	90.2%	80.0%	89.4%	90.3%	91.5%
	Fixed	89.9%	85.0%	89.1%	89.7%	91.4%

Note: This table shows the estimated medical loss ratio (100·costs/revenue) for each firm under each subsidy type. The 80% threshold is exceeded only in 2009 for one firm under price-linked equilibrium that is most favorable for the firms.

Therefore, for these two-insurer markets, the distortion from price-linked subsidies may be mitigated by MLR rules. In addition, other regulatory pressures (e.g., "rate review" by state insurance departments) may limit extremely high plan markups. These regulatory pressures are clearly necessary to explain prices in ACA counties with a single insurer, for which our model (taken literally) predicts infinite markups. For these extreme cases, our model predicts what insurers have an incentive to do – and may find ways of doing if they can elude regulatory barriers.

Table 11: Medical Loss Ratios with Two Insurers in 2011

		When Competing Against			
		CeltiCare	Network Health	BMC	NHP
MLR of:					
(Lowest cost)	CeltiCare		75%(81%)	78%(84%)	67%(74%)
↓	Network Health	91%(91%)		82%(87%)	75%(81%)
	BMC	89%(89%)	87%(87%)		75-76% (82%)
(Higher cost)	NHP	92%(91%)	91%(92%)	88-90%(91%)	

Note: This table shows the estimated medical loss ratio (100*costs/revenue) for the insurer of a given row, when competing only against the insurer of a given column. The first number in each cell is the MLR under price-linked subsidies; the number in parentheses is the MLR under fixed subsidies. The plans are listed in order of increasing costs: CeltiCare, Network Health, BMC, NHP. There are multiple equilibria under price-linked subsidies when only BMC and NHP are in the market, so there is a (small) range of MLRs.

D   Uncertainty simulations

To model the potential benefits of price-linked subsidies, we consider the problem of a regulator setting subsidies for an upcoming year. We assume that regulators observe lagged demand and cost data from insurers (e.g., from state insurance department rate filings), which give them all the relevant information for the optimal subsidy except for a *market-level* shock to cost growth.[9] This proportional shock applies to all plans' costs and the cost of charity care, multiplying them by a factor of $(1 + \Delta)$. Fixed subsidies have to be set in advance, prior to the realization of health care costs, but the regulator would like to set a higher subsidy when costs are higher because of the externalities of charity care and the desire to insure consumers against price variations (see discussion in Section I.B). If insurers observe cost shocks before setting prices, then prices contain information about optimal subsidies; price-linked subsidies can make use of this information.[10]

We assume the externality that is avoided when an individual chooses to purchase insurance has two components: charity care costs and the paternalistic component of the externality. Charity care costs are, unfortunately, quite difficult to measure. Instead of attempting to estimate them in our Massachusetts setting, we assume they are a proportion λ of an individual's expected costs in the average plan. For a cost-shock Δ,

$$C_i^{Charity} = \lambda \cdot \underbrace{\exp\left(\hat{\mu}X_i + \psi_0\right)}_{\text{Baseline cost in avg. plan}} \cdot (1 + \Delta) \qquad\qquad \lambda = \{.6, .8, 1\}.$$

Finkelstein, Hendren and Luttmer (2015) find that when individuals are uninsured, third parties cover about 60% of what their costs would have been if insured by Medicaid, implying $\lambda = .6$. The best case for price-linked subsidies, which we report in the text is $\lambda = 1$. We also show results for $\lambda = .8$ as an intermediate point.

We assume the regulator sets (income-specific) fixed subsidies to be optimal if costs are at the expected level (i.e., zero cost shock) and sets affordable amounts for price-linked subsidies to replicate the fixed subsidy amounts at $\Delta = 0$. In practice, we implement this by calibrating the paternalistic externality, E_0 as a residual to rationalize the ACA's affordable amounts (and therefore subsidy levels) as optimal at $\Delta = 0$. Our goal is to understand how cost uncertainty affects the case for price-linked subsidies assuming that subsidies would be set optimally if costs were known. If we do not calibrate E_0, our welfare results would be largely driven by the deviation of subsidies from their optimal level, absent a cost shock. Empirically, we find that a positive E_0 is needed to rationalize the ACA's affordable amounts. This result is consistent with the finding of Finkelstein, Hendren and Shepard (2019) that demand is less than marginal cost for most CommCare enrollees. With $E_0 = 0$ the implied

---

[9]It would also be possible to model insurer-specific cost shocks. However, market-level cost shocks are what matters for the welfare-relevant correlation between insurer costs and the externality of uncompensated care.

[10]We assume that consumers (like the regulator) do not observe the cost shock, so their utility of insurance (as reflected in ξ_j and β) is held constant. Even if consumers observed the cost shock, it is not obvious how their utility of insurance would be affected. Insurance against a larger risk is generally more valuable, but higher costs also mean the uninsured receive more charity care. Conceptually, if we allowed demand to shift with the cost shock, our simulated effects on coverage would change. But the basic welfare/policy implications – which are driven by the size of the subsidy and externality – would not clearly be different.

over-subsidizing of insurance would mean that the additional people insured under fixed subsidies would be a loss to the government, where really that loss is due to an affordable amount that is too high given the externality.

D.0.1   Prices and Coverage Rates

Figure 4a shows the price of the cheapest plan and the average subsidy for 2011 across values of the cost shock (on the x-axis) for each subsidy policy. Unsurprisingly, prices (shown in solid lines) increase with costs under both policies. They increase slightly faster under price-linked subsidies.[11] The figure illustrates how subsidies (shown in dashed lines) move quite differently under the two policies. Subsidies increase in tandem with the cost shock under price-linked subsidies, but are flat under the fixed policy (aside from small changes in the average, driven by changes in the income composition of insured consumers).[12]

Figure 4b shows the impacts on the share of eligible people who purchase insurance. Price-linked subsidies stabilize the insurance take-up rate (at about 40%) across cost shocks since they hold fixed consumer premiums for the cheapest plan, regardless of whether insurer prices rise or fall. However, under fixed subsidies coverage varies substantially: it is much higher (up to 70%) with a negative cost shock and lower (down to 19%) with a positive shock. Of course, it is not clear whether price-linked subsidies' stabilization of coverage in the face of cost shocks is desirable: it may be optimal for fewer people to buy insurance when it is more expensive.

D.0.2   Welfare

The three panels of Figure 5 show how public surplus changes with cost shocks under different assumptions on λ, which captures how the externality scales with costs. (The bottom panel is the same as Figure 3 in the text.) Fixed subsidies (in blue) are preferred at zero cost shock in all cases.[13] For non-zero cost shocks, the gap between fixed and price-linked subsidies narrows, particularly for larger values of λ. This occurs because the optimal subsidy gets farther from the fixed subsidy level as costs – and therefore the externality of uninsurance – diverge from expectations. If the externality increases one-for-one with costs ($\lambda = 1$), price-linked subsidies do better for cost shocks $\geq 15\%$ or less than -12.5%. However, as

---

[11]On average a 20% cost shock is about $75, so prices increase a little less than one-for-one with costs.

[12]We note that for the largest negative cost shocks that we simulate, prices under fixed subsidies fall slightly below the subsidy amount. In these cases, our calculations assume that insurers can charge "negative premiums" – i.e., rebate the difference to consumers. The ACA does not allow negative premiums, but implementing these does not seem infeasible. Rebates to consumers could be accomplished either via direct payments or via tax rebates (similar to the way the ACA's mandate penalty works). If rebates were not allowed, then fixed subsidies could also distort pricing incentives for highly negative cost shocks, since insurers would have no incentive to lower prices below the subsidy amount. This is not a major issue for the range of shocks we consider (where average subsidies are always less than the minimum price), but it could become important for even larger negative shocks.

[13]With the differing values of λ, the calibrated fixed externality E_0 differs, so the externality of infra-marginal consumers is not the same under the 3 scenarios, which is why the levels of public surplus do not line up across graphs. However, in each case, the units of public surplus are dollars per eligible person per month.

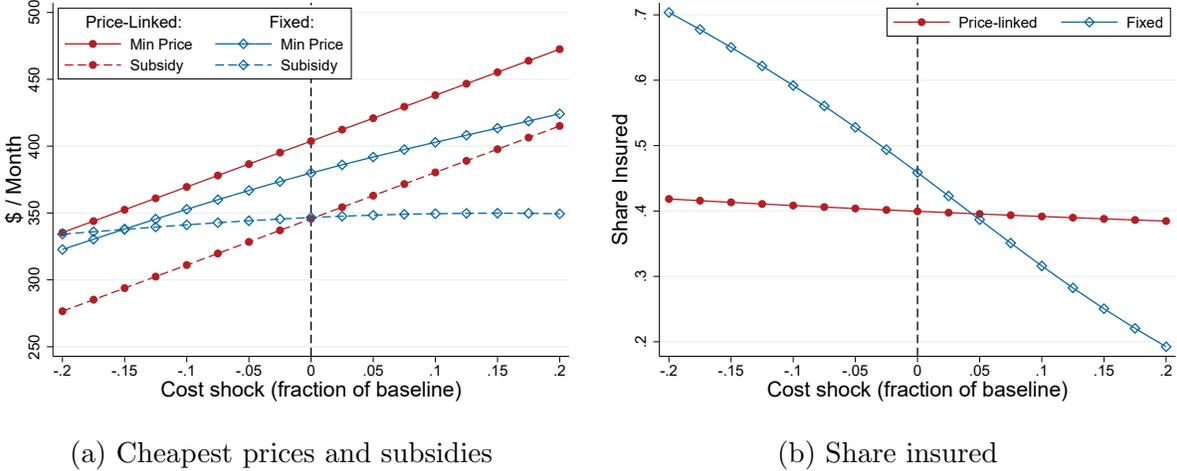(a) Cheapest prices and subsidies   (b) Share insured

Figure 4: Equilibrium under cost shocks

Note: These graphs show the equilibrium under price-linked and fixed subsidies for cost shocks of -20% to +20% of baseline. The left graph shows the price of the cheapest plan and the subsidy level. The right graph shows the share of consumers who purchase insurance.

the externality increases less with costs (lower λ), larger cost shocks are necessary to make price-linking better.

We then adjust the public surplus measure to include a cost of price variation to consumers, based on Equation (A.2). This adjustment depends on the level of consumption, which we approximate using monthly income at the middle of a consumer's income bin, minus the average premium (or mandate penalty) they pay. For the priced difference (ΔP), we use the premium variation for the average plan relative to its premium at $\Delta = 0$. The cost of price variation also depends on the coefficient of relative risk aversion, γ. Chetty (2006) argues that $\gamma \leq 2$; to consider the case most favorable to price-linked subsidies, we use $\gamma = 2$ (dashed lines). We also consider a more extreme case of $\gamma = 5$ (dotted lines), to capture the idea that society might have a strong concern about "affordability" or to proxy for factors like consumption commitments that increase local risk aversion (Chetty and Szeidl, 2007). For both values of γ, the welfare losses from greater price risk to consumers are fairly small and do not substantially change our results – i.e., the dashed and dotted lines in 5 are not very different from the solid lines. This is explained by the well-known fact that for reasonable risk aversion, utility is locally quite close to linear (Rabin, 2000). The maximum premium variation we consider (about +/-\$75 per month) represents only about +/-5% of average income, even for this low-income population. Based on this analysis, it seems that insuring consumers against price risk is a less important rationale for price-linked subsidies than the correlation between prices and the externality of charity care.

Results for 2009

Figure 6 shows public surplus under fixed subsidies and under the minimum price-linked subsidy equilibrium in 2009, for cost shocks of different sizes. The results are qualitatively the same as 2011, but since the initial difference between the two subsidies is smaller, a
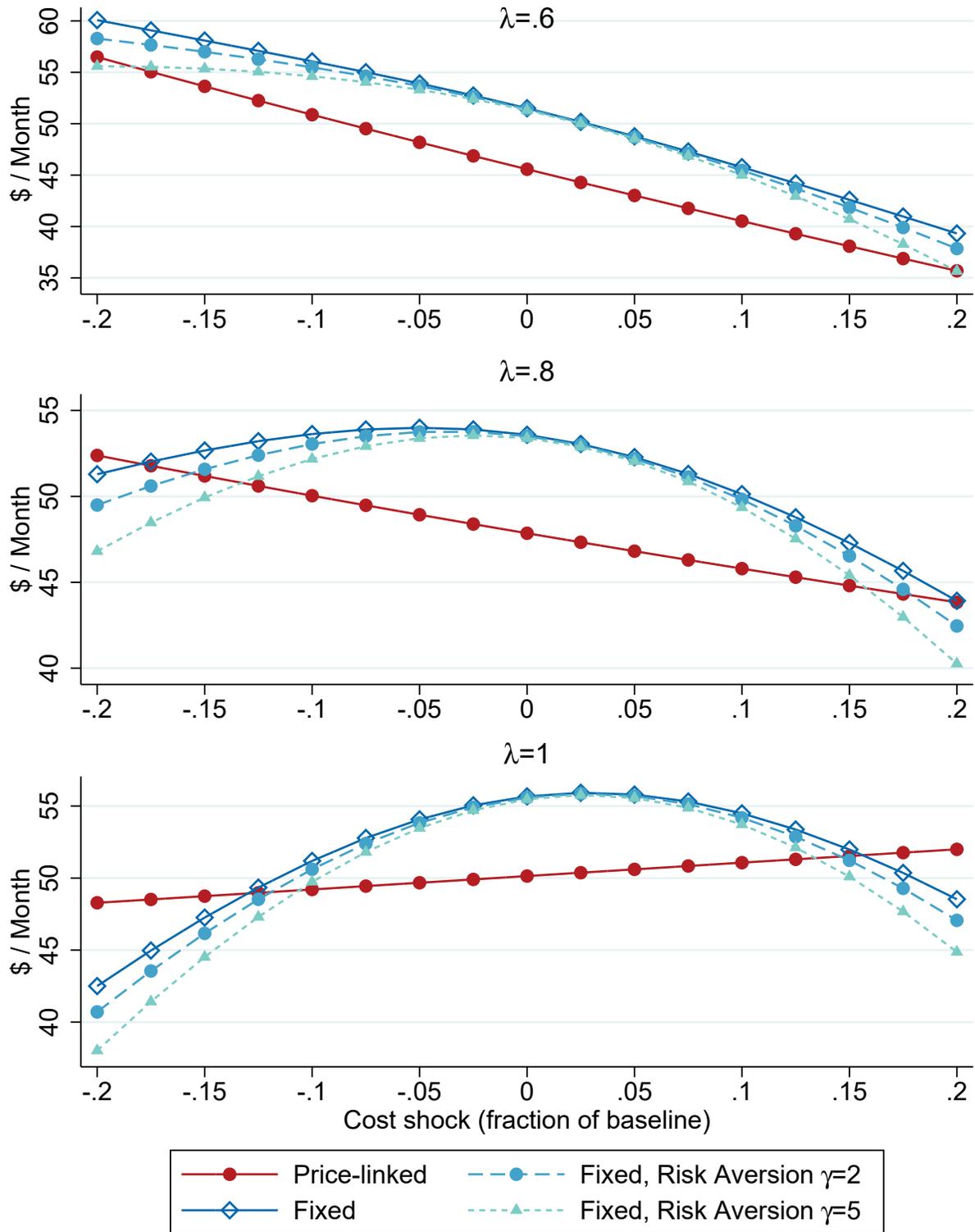
Figure 5: Public surplus under cost shocks in 2011

Note: These graphs show public surplus (in dollars per month per eligible member) under price-linked and fixed subsidies for cost shocks of -20% to +20% of baseline. Each graph corresponds to a different assumption about how much the externality changes with costs. The dashed and dotted lines subtract a cost of pricing risk to consumers (using the method in Equation (A.2)) with coefficients of relative risk aversion of 2 and 5, respectively.

smaller cost shock is required to make price-link subsidies better than fixed subsidies – typically a shock of between 5-10%.

Including insurer profits in welfare

Regulators may also include profits in their objective function. We repeat the comparison of fixed and price-linked subsidies under uncertainty for this more inclusive welfare definition. Doing so requires recalibrating the fixed component of the externality of uninsurance, E_0, since it is calculated based on the assumption that equilibrium subsidies are optimal and what is optimal depends on the objective function. In practice, we find that the calibrated E_0 is often negative. This occurs because if the regulator cares about profits, they will want to subsidize purchases even if there is no externality because price (willingness to pay) is above cost, so there are social gains to the marginal purchase. Although a negative E_0 seems non-intuitive, the uncertainty analysis only makes sense if baseline subsidies are optimal at zero cost shock, so we proceed with the estimated values.

Figure 7 shows welfare including profits under fixed and price-linked subsidies for different cost shocks. Fixed subsidies do not do quite as well as without profits, but for $\lambda = .8$ fixed subsidies are better than price-linked for cost shocks $\leq 12.5\%$ and $\geq -10\%$.
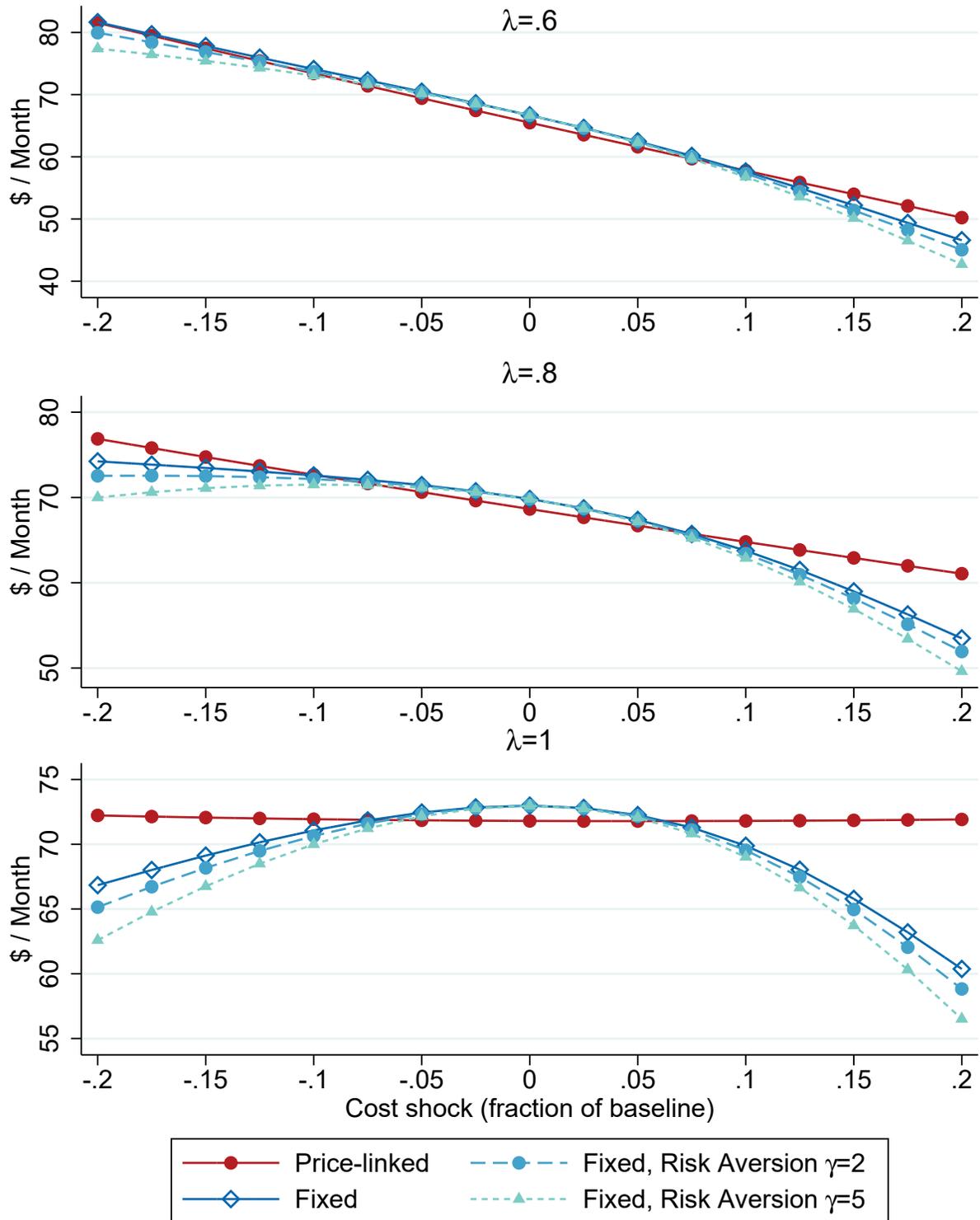
Figure 6: Public surplus under cost shocks in 2009

Note: These graphs show public surplus under price-linked and fixed subsidies for cost shocks of -20% to +20% of baseline. Each graph corresponds to a different assumption about how much the externality changes with costs. The dashed and dotted lines are risk adjusted with factors of relative risk aversion of 2 and 5, respectively.
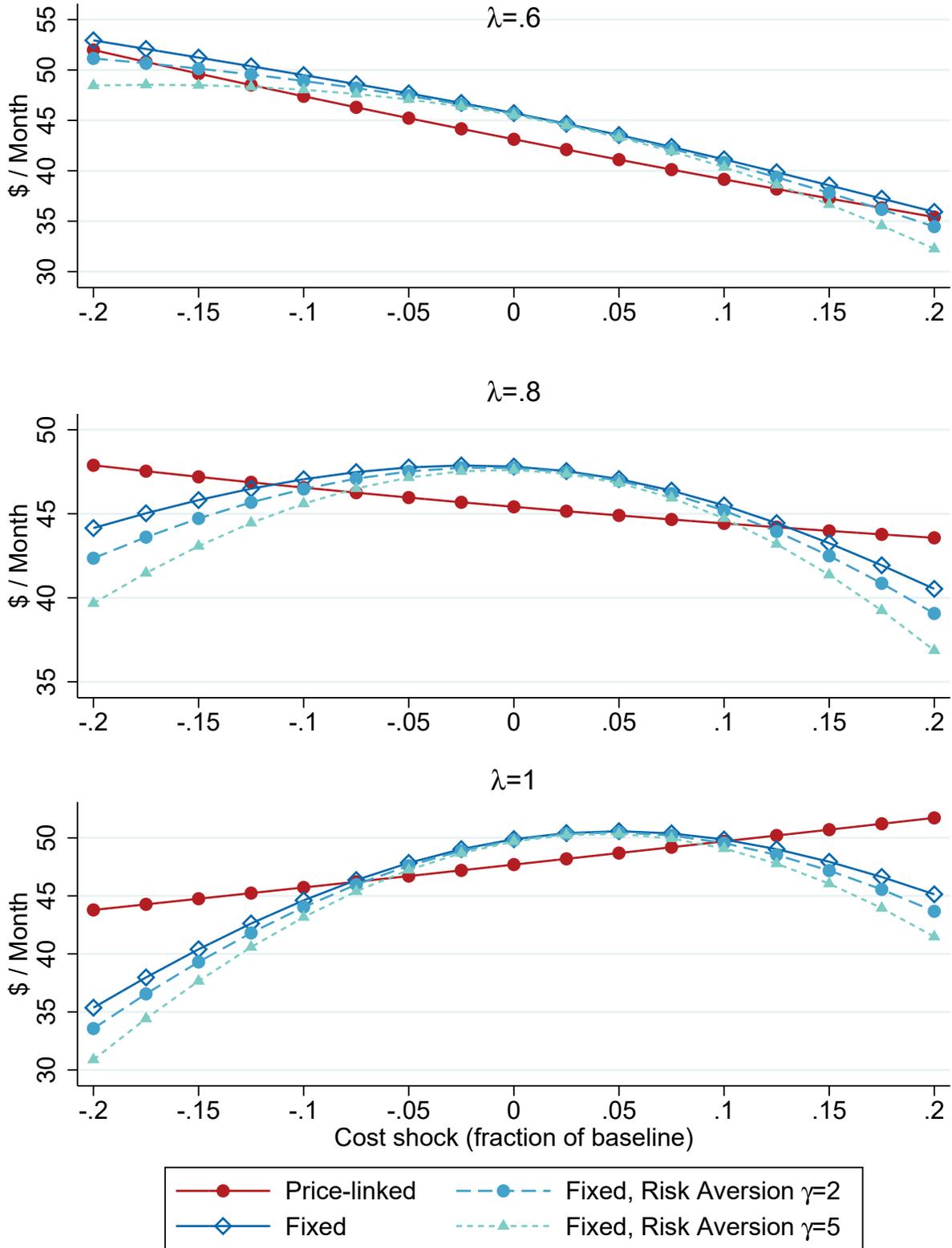
Figure 7: Public surplus plus profits under cost shocks in 2011

Note: These graphs show the sum of public surplus and profits under price-linked and fixed subsidies for cost shocks of -20% to +20% of baseline. Each graph corresponds to a different assumption about how much the externality changes with costs. The dashed and dotted lines are risk adjusted with factors of relative risk aversion of 2 and 5, respectively.

31

References

Berry, Steven, James Levinsohn, and Ariel Pakes. 2004. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." *Journal of Political Economy*, 112(1): 68.

Chetty, Raj. 2006. "A new method of estimating risk aversion." *The American Economic Review*, 96(5): 1821–1834.

Chetty, Raj, and Adam Szeidl. 2007. "Consumption Commitments and Risk Preferences." *Quarterly Journal of Economics*, 122(2).

Commonwealth Care. 2008. "Report to the Massachusetts Legislature: Implementation of the Health Care Reform Law, Chapter 58, 2006-2008."

Finkelstein, Amy, Nathaniel Hendren, and Erzo FP Luttmer. 2015. "The Value of Medicaid: Interpreting Results from the Oregon Health Insurance Experiment."

Finkelstein, Amy, Nathaniel Hendren, and Mark Shepard. 2019. "Subsidizing health insurance for low-income adults: Evidence from Massachusetts." *American Economic Review*, 109(4): 1530–67.

Rabin, Matthew. 2000. "Risk aversion and expected-utility theory: A calibration theorem." *Econometrica*, 68(5): 1281–1292.