Online Appendix

"Estimating welfare in insurance markets using variation in prices" by Einav, Finkelstein, and Cullen

A.1. Data construction

Construction of the baseline sample. We make a number of sample restrictions. First, we make a number of restrictions for purposes of data purity, which brings the original sample of about 45,000 active employees down to about 37,000 active employees. The biggest reduction in sample size comes from excluding employees who are not at the company for the entire year (for whom we do not observe complete annual medical expenditures, which are necessary for estimating the cost curve). In addition, we exclude employees who are outside the traditional benefit structure of the company (for example, because they were working for a recently acquired company with a different (grandfathered) benefit structure). For such employees we do not have detailed information on their insurance options and choices. We also exclude a small number of employees because of missing data or data discrepancies.

Second, because the new set of health insurance options we study did not apply to many hourly employees and because (as we discuss in Section IV.B) the pricing variation is cleaner for the salaried employees, we further limit the analysis in this paper to salaried employees, who are approximately one third of the U.S.-based Alcoa workforce.

Third, to illustrate most easily how the theoretical framework maps to the empirical strategy, we limit the baseline analysis to the two modal health insurance choices: a higher and a lower level of PPO coverage, to which we refer as contract H and contract L throughout Section IV of the paper. Approximately two-thirds of salaried employees chose one of these two PPO options. In Section A.2 of this online appendix we describe the other options in more detail and show that the pricing of the two PPO options we focus on does not affect the probability of the employee choosing one of the other options we do not analyze. This helps to alleviate concerns about potential biases from our sample selection on contract choice.

Finally, for simplicity, our baseline specification further limits our sample of salaried employees who choose either contract H or contract L to the slightly over one half of such employees who choose the most common coverage tier, which is family coverage. All employees have a choice of four different tiers for health insurance coverage: employee only, employee plus spouse, employee plus children, and family coverage. In Section A.2 of this online appendix we show that our results are similar when we include employees in all coverage tiers. We assume throughout that the choice of coverage tier is unrelated to the pricing variation. A priori, this seems a reasonable assumption given that coverage tier options are limited by the demographic composition of the family, and that the price multiplier across coverage tiers is the same for all employees.[1] Consistent with our

[1] Specifically, for any health insurance coverage option, for all employees the family price is always triple the

assumption, we find that the (relative) price of contract H in the family coverage tier does not predict (either economically or statistically) which coverage tier the employee chooses (not shown).

Table A1 provides some descriptive statistics on the employees. Column (1) presents descriptive statistics for the sample of 37,000 active employees for whom we have complete data. Column (2) limits the sample to the approximately one third of the sample who are salaried employees. Column (3) makes the further (minimal) restriction to the salaried employees who face the new benefit design. Column (4) further limits the sample to employees who choose either contract H or contract L, and column (5) further limits the sample to those in family coverage. Column (5) represents our baseline sample that we use for most of the empirical analysis. Section A.2 of this online appendix presents analyses that use all coverage tiers (column (4)) and all coverage options (column (3)).

For comparison, columns (6) through (8) of Table A1 present statistics from the 2005 March Current Population Survey (CPS) on characteristics of various types of full-time employees in the U.S. The principal (and unsurprising) finding is that Alcoa employees do not appear to be representative of any cut of full time employees in the U.S. We also compared the medical expenditures in our baseline sample to medical expenditure data from the 2004 Medical Expenditure Panel Survey (MEPS). Salaried employees in Alcoa tend to have about 50 percent lower medical expenditures than comparable individuals in the MEPS.[2] This may be because Alcoa salaried employees are healthier than the general population or that they tend to live in regions with lower healthcare costs.[3] Such comparisons underscore our statements that our empirical results should be viewed as an illustrative example of how our proposed approach can be applied, rather than as generalizable findings about the employer-provided health insurance market in the U.S.

Construction of incremental costs. One of the key variables in our analysis is the insurer's incremental costs c_i. This is defined as $c(m_i; H) - c(m_i; L)$, where $c(m_i; j)$ is the cost to the insurer from covering medical expenditures m_i under contract j. We note that medical utilization (m_i) is held fixed in the construction of c_i, so there is no estimation involved in the process.

The construction of c_i requires detailed knowledge of each plan's benefits as well as individuals' realized medical claims. We obtained the former from reading each plan coverage details and verifying them with the actual reimbursements we observe in the data. The latter is part of our data, which include detailed information about every single claim made by Alcoa employees during 2004. For each claim we know the claim date, the claim amount, how much of it was reimbursed by Alcoa, and how much was paid out of the insured's pocket. For the latter we also know whether it was applied to the annual deductible or was part of a coinsurance. We also know whether each

"employee only" price, 1.58 times the "employee plus children" price, and 1.43 times the "employee plus spouse" price.

[2] Specifically we focus on MEPS observations on individuals with full-year coverage by employer-provided health insurance, and we try to reweight observations to adjust for age and gender.

[3] In addition to the non-representativeness of health expenditures in our Alcoa population, we further note that our cost variable is a complicated non-linear transformation of total cost, which is perhaps even more context-specific (as the transformation depends on the particular features of the plans we study).

claim was associated with in-network or out-of-network care, and additional medical details which are less relevant for the construction of c_i.

To construct c_i we simply "run" the observed set of claims for each covered employee (and his dependents) through the reimbursement rules twice. Once by applying the rules of contract H and once by applying the rules of contract L. A key feature of our setting which facilitates this construction is that the two contracts we focus on vary only in their employee cost-sharing rules. Alcoa is the direct insurer of both plans, and the plans are identical in all other features, such as the network definition and the benefits covered. As a result, we do not have to worry about differences between contracts H and L in plan features that might differ in unobservable ways across employees (for example, differences in providers or provider prices, the relative network quality, and so forth). In particular, this implies that a set of claims submitted under one contract would be eligible (and identical) claims under the other contract. Once we have "run" the claims for each employee for each contract, we have obtained $c(m_i; H)$ and $c(m_i; L)$, and the difference is our constructed variable c_i.

Applying the plans' rules is fairly simple, although certain issues require some care. One such issue is whether a claim was made in network or out of network, since different deductible and cost sharing rules would apply (see Figure III). A second issue is related to preventive care. Alcoa provides full coverage (with zero out-of-pocket payments) for various preventive treatments, including periodical exams, well baby, etc. It is therefore important to know whether claims are associated with preventive-related services, since cost sharing rules do not apply to such claims. A third issue, which is typical of most health insurance plans, is the interaction between an individual deductible and a family deductible (as well as analogous issues regarding individual and family out-of-pocket maximums). In our data, the family deductible is always twice the individual deductible. For a family with more than two covered individuals, it is therefore important to account for the interaction among family members, as the cost sharing rules would vary depending on how the spending is distributed among the family members. That is, a given individual in a family can exhaust his deductible either by spending his individual deductible or by having the cumulative spending of other members of the family reach the family deductible. In the construction of c_i, we therefore need to account for the composition of spending within the family.

Fortunately, the data are quite detailed and the plan rules are fairly simple (despite the above issues), allowing us to calculate $c(m_i; j)$ with a great deal of accuracy. Indeed, our calculated reimbursements (based on our application of the plan rules) and the actual reimbursements observed in the data are almost the same. For example, for individuals with contract H the correlation between their actual (observed) share of out-of-pocket spending (out of total expenditure) and our constructed share is over 0.97. The same is true for contract L, or when we correlate levels of expenditures instead of shares. Recall that we still need to apply our construction, because for each individual we only observe the actual reimbursement for the contract he chose, while the second element of c_i is always a counterfactual. For consistency, we never use the actual reimbursement and always compute c_i by constructing both elements, $c(m_i; H)$ and $c(m_i; L)$.

A.2. Robustness of the baseline estimates

In this section we explore the sensitivity of our welfare estimates to a number of alternative specifications. Our overall finding is that the magnitude of the various welfare estimates discussed in the paper – even those that involve extrapolation considerably out of sample – are qualitatively similar across a range of alternative specifications. In particular, across various alternative specifications, the welfare gain from a price subsidy that achieves the efficient price is always substantially below the social cost of the required price subsidy. Similarly, the welfare loss from competitive pricing when choice over contracts is allowed is always lower than the welfare loss from mandatory coverage by contract H, and the welfare cost of competitive pricing is always less than 10 percent of the total surplus that could be generated from efficient pricing. In the end of this section we also address possible concerns regarding sample selection.

Functional form and theoretical restrictions on the demand curve. Table A2 summarizes some of the sensitivity analyses. Panel A summarizes the implied welfare implications of each specification. For completeness, Panel B shows the corresponding parameter estimates from each specification (which are used to derive the welfare estimates shown in Panel A). In the interest of brevity we focus our discussion primarily on the robustness of the resultant welfare estimates (columns (6) through (8) of Panel A), which are our main interest. The first row of Table A2 presents the results from our baseline specification reported in the main text (see Table III). Subsequent rows report results from a single, specified departure from this baseline.

Rows 2-5 in Table A2 explore the sensitivity of our results to our functional form assumptions. Row 2 shows the results from our baseline specification are quite similar if we estimate a probit for the demand equation rather than a linear demand. Unrestricted quadratic demand (not reported) behaves very badly out of sample and is therefore not shown (but in row 5 we report and discuss a restricted specification that includes a quadratic demand curve). As can be seen in Figure V, the linear specification fits the cost data well.[4]

We also experimented with imposing restrictions on the demand curve that are implied by basic price theory. Willingness to pay is (theoretically) bounded from above at $1,500 (the maximum possible out-of-pocket savings from contract H; see Figure III) and (theoretically) bounded from below by 0 (any rational individual should always prefer more coverage to less if the former is offered for free). Our baseline demand estimate (Table A2, row 1) satisfies the first constraint (the share of contract H becomes 0 at a price of $1,350), but not the second. At a price of 0, the share of contract H is only 0.94.[5] The results in row 3 show that constraining the share of contract H to be 1 when price is 0 does not noticeably affect our welfare estimates. Row 4 shows the results are

[4] We explored alternative functional forms for the cost curve, such as a quadratic, log-log, and log-linear functions. Not surprisingly, the results (not shown) were very similar in sample. However, curvature (concavity in particular) in the estimated AC curve sometimes led to out-of-sample predictions that were difficult to interpret (such as non-monotone MC curve). Given all these hard-to-interpret predictions were driven by out-of-sample predictions from an ad hoc functional-form extrapolation, we prefer to simply reject such extrapolations and focus our discussion on those extrapolation that seem to "better behave" (out of sample).

[5] One reason why we may estimate demand below 1 for a price of 0 is that our functional form assumption of

also similar if we impose the constraint that willingness to pay is bounded at $800, which may be a more reasonable upper bound in practice than the theoretically possible $1,500.[6] Row 5 estimates a quadratic demand curve, imposing both the (1,$0) and the (0,$800) constraints on (Q,P), and again resulting in welfare estimates that are quite stable.

Tax treatment of employee premiums. We also considered the sensitivity of our results to the tax treatment of employee contributions to health insurance and to out-of-pocket medical expenditures. Employee premium contributions are made pre-tax. Employees can pay their out-of-pocket medical spending pre-tax as well, by contributing to a Flexible Spending Account (FSA). If all out-of-pocket expenses were paid pre-tax, the tax treatment of employee premiums and employee medical spending would be symmetric, and ignoring the tax subsidy to employee premiums (as we do in our baseline specification) would be appropriate. However, in practice, less than a quarter of Alcoa employees contribute to an FSA. It is of course unclear whether employees who do not take advantage of the tax subsidy to out-of-pocket medical spending offered by FSAs are cognizant of the tax subsidy to employee premiums. However, to investigate the sensitivity of our findings to the tax subsidy, we consider the effect on our estimates of assuming that all employees (including those who contribute to FSAs) make their health insurance choices based on the pre-tax price.[7] We calculate the average tax subsidy (i.e. one minus the average marginal tax rate) for our sample to be 65 percent.[8] In row 6 we therefore re-estimate the baseline specification with the price variable in both the demand and cost equations multiplied by 0.65. Once again the core welfare estimates are

---

linear demand is not appropriate for extrapolating this far out of sample. Another possible explanation may be that contract L was the default option in 2004. We suspect that default may be less important in our setting than in others because 2004 was the first year in which the new benefits were offered. These new benefits came with much effort by Alcoa to advertise and explain the new options to its employees, making it likely that most individuals were "active" choosers. Moreover, it is possible to have a model of defaults in which our welfare anlaysis is unaffected. We discuss this in a little more detail below.

[6] $1,500 out-of-pocket savings from contract H is only possible if the covered family members spend enough in-network *and* out-of-network to hit the (separate) out-of-pocket maximums. In practice, this never occurs. Indeed, none of the employees in our sample hits the out-of-pocket maximum out-of-network and only about 1 percent hits the in-network out-of-pocket maximum. A potentially more reasonable constraint therefore is that willingness to pay for contract H should not exceed $800, which is the reduction in out-of-pocket expenditures associated with contract H if the family spends more than the deductible in-network and more than the deductible out-of-network but less than the amount that would cause them to hit the out-of-pocket maximum (see Figure III).

[7] We do not observe in the data which individuals participate in the FSA.

[8] The tax subsidy is given by $(1 - \tau_f - \tau_s - \tau_{ss} - \tau_{mcr})$ where τ_f is the federal marginal tax rate, τ_S is the state marginal tax rate, τ_{ss} is the marginal Social Security (FICA) payroll tax on the employee, and τ_{mcr} is the marginal Medicare payroll tax on the employee. We estimate these marginal tax rates using the NBER's TAXSIM model, which takes as inputs the major determinants of marginal tax rates and computes the various marginal rates just mentioned. Many of the required data elements (or reasonable proxies for them) are available in our company's data, including annual wage and salary income, state, marital status, number of dependents and ages of family members. We assume all employees with family coverage file jointly and do not itemize. We impute wage and salary income of spouse, property income, and dividend income based on the ratio of each of these variables to own income for the sample of full time, white collar manufacturing employees in the March CPS; we pool the 2004-2007 March CPS to increase sample size (Table A1, column (8) presents descriptive statistics for this sample in the March 2005 CPS). All other inputs required by TAXSIM are assumed to be zero. For more information on TAXSIM, see www.nber.org/taxsim.

not noticeably affected, although naturally our estimates of the equilibrium and efficient allocations (see columns (1) through (4)) shift considerably.

Additional covariates and alternative samples. Our baseline estimates of the demand and cost curves include no covariates in the analysis besides the (relative) price. Only variables that are priced should be controlled for in our analysis of selection and its welfare costs. The fact that, for example, individuals of, say, different incomes or different ages may have different expected medical costs, and that this may affect which plan they choose, is part of the endogenous selection we wish to study, rather than control for, since these characteristics are not priced. However, to allow for the possibility that the price menu may be selected differently across states in a systematic fashion (e.g., reflecting differences in healthcare costs across states), in row 7 we include state fixed effects in the demand and cost estimates. Although our estimates become somewhat less precise (see Panel B of Table A2), the welfare implications remain quantitatively similar (Panel A). In row 8 we add all of the contemporary employee characteristics (see Table I) as covariates to the demand and cost curves (in addition to the state fixed effects).[9] Once again the results are similar. The fact that the slope of the estimated demand curve remains similar is unsurprising given the evidence in Table I that pricing is orthogonal to these employee characteristics. The fact that the slope of the estimated cost curve remains similar suggests that the adverse selection we detect is not driven by the fact that in our setting the observable characteristics of employees are not priced.[10]

Finally, in row 9 we estimate our baseline specification using all four coverage tiers rather than just employees with family coverage. Since prices vary by coverage tier, we include (de-meaned) indicator variables for the coverage tier in both the demand and cost estimates.[11] The parameter estimates and welfare implications are quite similar to our baseline results.

We also tried restricting our baseline sample, specifically by excluding the 199 individuals who face the $570 (relative) price, which seem likely to affect the demand estimates. Indeed, we found that eliminating this points substantially reduces the demand elasticity (by about 45%) and it is no longer statistically significant. However, when we do so the average cost curve remains similar. Thus, the steeper demand curve produces a steeper marginal cost curve, exacerbating the welfare costs of inefficient pricing due to selection. As a result, despite the steeper demand curve (which all else equal should reduce welfare costs), our welfare estimate remains roughly the same ($9.77 compared to $9.55 in the baseline specification). This type of robustness exercise illustrates that it is the combination of the demand and cost curves that together contribute to the magnitude of

[9] In both rows 7 and 8 the covariates are demeaned so that the constant term is comparable across specifications.

[10] In this sense, the robustness test on the cost curve is one sided. Had we found that the slope of the cost curve changed once we controlled for unpriced observables, this would not necessarily be a cause for concern. It could simply reflect the fact that much of the selection in our setting is driven by these unpriced observables.

[11] The price variable is defined for the chosen coverage tier. As noted earlier, for all employees the prices of contracts in the other coverage tiers are always the same fixed multiplier of the prices in the family coverage tier. To account for the fact that for "employee only" coverage the deductible and out-of-pocket maximum is half of what it is for the other three coverage tiers, we multiply price (p_i) and cost (c_i) by two for the 16% of employees with "employee only" coverage.

the welfare loss.

Possible sample selection. An important potential concern with all of the foregoing analyses is that we limit the sample to only those who choose contract H or contract L, and exclude the approximately one-third of salaried employees who chose one of the five other available options. These five other options are an HMO (chosen by about 7% of salaried employees), opting out of any employer-provided coverage (about 8%), two even lower coverage PPO options (3% in the two of them combined), and a Health Reimbursement Account (HRA) PPO option, which combines a high deductible health insurance policy with tax preferred employer contributions that can be used to pay out-of-pocket expenses (approximately 17%).[12]

In practice, however, our analysis suggests that our sample selection is unlikely to have important effects on our demand estimates (and, of course, it is irrelevant for the estimate of the cost curve which by design is run on the endogenously selected sample of individuals choosing contract H). In particular, we found that the price of contract H relative to contract L (our key right-hand-side variable) does not predict whether or not the employee "opts in" to one of the two contracts we study (contract H and contract L), as opposed to "opting out" into one of the remaining options. We suspect that this in part reflects the fact that many of the other options (in particular the three with non-trivial market share, the HMO, opting out of insurance, and the HRA) are quite horizontally differentiated.

Table A3 presents some of these findings. The dependent variable in the reported linear regressions is an indicator variable that takes the value of 1 if the employee chose one of the "outside goods" and 0 if he chose either contract H or contract L. The right-hand-side variable p is (as before) the relative price of contract H compared to contract L. Column (1) reports the results for employees with family coverage. We find that a \$100 increase in the (relative) price of contract H is associated with an economically and statistically insignificant decline (of 0.09 percentage points) in the probability of choosing one of the outside goods. Column (2) shows similar results when all coverage tiers are pooled. A complication with both of these analyses is that because coverage tier is not available for the 8% of the sample who opt out of coverage, these employees are excluded from the analysis. In column (3) therefore we include in the sample the employees who opt out of coverage. However, since coverage tier is not known for these employees we cannot control for coverage tier and, moreover, we can no longer define the price variable based on the coverage tier. We instead assign all employees the family prices regardless of what coverage tier they actually chose (if known).[13] Once again there is no evidence that the relative price of contract H has an

[12] The in-network deductibles for the two lower coverage PPO options are \$1,000 and \$1,500. The high deductible HRA PPO has a \$3,000 in-network deductible, but the employee receives \$1,250 tax free from Alcoa each year which can be spent on eligible medical expenses (including but not limited to the plan's cost sharing provisions). Unspent funds in the HRA can be rolled over the subsequent years, but any unspent balance is forfeited upon separation from Alcoa. The out-of-pocket maximums of all these options are also higher, but this is largely irrelevant empirically since it is extremely rare (less than 1%) for any employee to hit the out-of-pocket maximum of even the most generous option. Coinsurance rates are the same across all PPOs.

[13] Since, as noted, the prices of other coverage tiers are proportional to the family price, this is not an unreasonable

economically or statistically significant effect on the probability of choosing the outside good.

A.3. Extensions

In this section we briefly discuss several extensions to our application, following the discussion of possible extensions to our framework in Sections II and III.

More than two coverage choices. As noted in Section III, it would be conceptually straightforward to extend our empirical analysis to consider more than two choices. However, we face practical obstacles to doing so in our setting. In particular, as is typical in data sets like ours, we do not observe medical expenditures for employees covered by an HMO or who opted out of employer-provided coverage. We therefore cannot estimate the cost curve for these options. It is also difficult to model the demand for these two options, since the prices are not known, nor is it entirely clear how to define the "good" being purchased.[14] We experimented with estimating demand and cost systems for the remaining five PPO options. However, the relatively small sample sizes on the other three PPO options combined with the relatively high multi-collinearity in relative prices among the different PPO options resulted in fairly imprecise (and therefore relatively uninformative) estimates of the demand and cost systems.

Moral hazard. As we discussed in Section III.B, our framework also allows us to easily test for and quantify moral hazard, which is defined by the vertical distance between MC^H and MC^L. Moreover, as discussed in Section II.D, when contract L provides partial coverage (as in our application) moral hazard will affect the welfare analysis. Therefore it is important to examine moral hazard empirically in our setting.

With two partial coverage contracts, c_i^H is defined as the incremental cost to the insurer of covering employee i with contract H rather than with contract L assuming i behaves as if he is covered by contract H. Analogously, c_i^L is the incremental cost to the insurer of covering employee i with contract H rather than with contract L assuming i behaves as if he is covered by contract L. Our foregoing estimates of AC, which were estimated on the sample of individuals who chose contract H, therefore gives us AC^H. And our estimate of MC, using our estimate of AC^H and our estimate of the demand curve for H (equation (11)), similarly gives us MC^H. To estimate AC^L we estimate the same cost equation (equation (12)) but on the sample of individuals who chose contract L. To back out MC^L from AC^L we use the demand curve for contract L, i.e. equation (11) estimated with D_i replaced by $1 - D_i$.

We have run this exercise on our baseline sample and were unable to reject the null of no moral hazard (i.e. $H_0 : MC^L = MC^H$). Our estimates were quite imprecise, suggesting that we may lack

approach.

[14] The price of the HMO is literally not known, and likely varies across geographic areas. Employees receive a \$1,000 "credit" if they opt out of any coverage. However, without knowing what price they face for purchasing insurance outside the company it is not clear what the true price is. Relatedly, in contrast to the PPO options, the characteristics of the HMO option and any coverage offered outside the firm are not known.

sufficient power in our setting to detect moral hazard. This may not be surprising given that the design of the insurance contracts in our setting (see Figure III) should make moral hazard primarily affect those employees who expect to spend less than the contract L deductible. In practice, this is likely to be a small fraction of our data.[15]

As a different way to make this point, we applied the widely used moral hazard estimate of Manning et al. (1987)[16] from the Rand Health Insurance Experiment to the total spending of each employee covered by contract H. We assumed a price effect which is based on the change in the marginal cost-sharing this employee would face under contract L compared to contract H, holding his realized (rather than expected) spending fixed. This back-of-the-envelope calculation led to an average change in insurer's cost of 3%, driven by the fact that three quarters of the employees did not experience any change in marginal cost sharing. In light of this, we find it unsurprising that it is hard to detect moral hazard in this setting.

Departures from revealed preference. As we noted at the outset, our approach to welfare analysis has relied on revealed preferences. It is possible to use our framework for welfare analysis when we are not willing to assume revealed preferences, although this would require specification of the precise alternative choice model and how it maps to welfare. Some "behavioral" models are easily translated to our approach. Consider, for example, the possible role of defaults. The default option in our setting is contract L. If one believes that there is a (constant) fraction α of the sample who always chooses the default, then it is possible to implement our approach, and perform welfare analysis on the remaining $1 - \alpha$ share of the sample, who are "active" choosers.

[15] Considering in-network spending, there are 9% of the employees in our baseline sample who spend less than the contract L (in-network) deductible of $500. Out-of-network spending would increase this share (but not by much).

[16] Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, Arleen Leibowitz, and M. Susan Marquis, "Health insurance and the demand for medical care: Evidence from a randomized experiment," *American Economic Review,* 77 (1987), 251-277.

Table A1: Summary statistics

	2004 Company Data					March 2005 CPS		
	All employees	Only salaried employees	Only salaried employees with new benefit design	Col. (3) limited to only employees who chose *H* or *L*	Col. (4) limited to employees with family coverage	All full-time employees	Only in manufacturing	White collar employees in manufacturing
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Individuals	36,814	11,964	11,325	7,263	3,779	83,118	11,178	4,688
Fraction Male	0.78	0.73	0.73	0.77	0.86	0.58	0.70	0.64
Fraction White	0.77	0.87	0.86	0.86	0.86	0.82	0.82	0.86
Fraction unionized	0.33	0.00	0.00	0.00	0.00	0.13	0.14	0.04
Age								
Mean	44.24	44.51	44.50	45.17	42.66	41.39	42.13	42.87
Std. Deviation	9.86	9.22	9.21	9.12	7.22	12.33	11.45	10.88
Median	45	45	45	46	43	41	42	43
Tenure with company (years)								
Mean	13.23	13.26	13.23	13.69	12.70	n/a	n/a	n/a
Std. Deviation	10.28	9.95	9.96	10.01	8.93	n/a	n/a	n/a
Median	11	12	12	13	12	n/a	n/a	n/a
Annual Salary (current $US)								
Mean	53,103	71,622	72,821	74,017	80,999	41,869	46,195	63,157
Std. Deviation	47,642	77,936	79,373	91,530	112,790	47,955	45,435	58,072
Median	47,283	60,484	61,433	61,822	66,335	32,000	35,000	50,000

Columns (1) to (5) present summary statistics for different cuts of the 2004 Alcoa employees. Column (1) presents statistics for all active employees in our sample, column (2) for salaried employees only. Column (3) looks at a slightly smaller group of salaried employees who faced the new benefit design, and column (4) further restricts attention to salaried employees who chose either contract H or contract L (who are the primary focus of our analysis). Column (5) further limits the analysis to those who chose family coverage; this sample is used to generate our baseline estimates. For comparison, columns (6) to (8) present summary statistics for full time employees (defined as those who on average worked 35 or more hours per week in the previous year) in the March 2005 CPS. Column (6) shows all full time employees, column (7) shows all full time employees in manufacturing industries, and column (8) shows all full time white collar employees (defined based on occupation codes) in manufacturing industries; in these three columns we use CPS sampling weights ("earning weights" for the union variable, and "person weights" for all others).

Table A2: Robustness

Panel A: Welfare estimates from different specifications

	Competitive Equilibrium		Efficient Allocation		Welfare cost of Adverse Selection			
	Q	P	Q	P	In dollar terms (per market participant)[a]	Relative to social cost of efficient subsidy[b]	Relative to welfare cost of mandating contract *H*[c]	Relative to total achievable welfare[d]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1 Baseline (family coverage, no state fixed-effects)	0.617	463.51	0.756	263.94	9.55	21.1%	32.4%	3.4%
Robustness to demand estimates								
2 Probit demand	0.619	463.59	0.790	187.85	11.32	17.3%	31.6%	3.5%
3 Linear demand, constrained to go through (Q,P)=(1,$0)	0.612	463.56	0.750	299.04	7.81	21.1%	30.2%	3.4%
4 Linear demand, constrained to go through (Q,P)=(0,$800)	0.562	463.59	0.688	387.90	3.30	21.1%	16.4%	3.4%
5 Quadratic demand, constrained to go through (1,$0) and (0,$800)	0.587	463.58	0.738	343.51	5.00	18.8%	45.6%	4.5%
Robustness to tax subsidy								
6 Baseline specification, but accouting for pre-tax premiums	0.389	514.49	0.567	348.53	7.71	27.3%	16.8%	9.8%
Robustness to sample and source of variation								
7 State fixed-effects included (in both demand and cost regressions)	0.622	460.16	0.699	341.40	3.65	14.6%	6.5%	1.2%
8 State fixed-effects and demographics included (in both regressions)	0.641	440.00	0.724	306.67	4.42	15.3%	9.2%	1.3%
9 All coverage tiers, no state fixed-effects[e]	0.593	434.20	0.704	244.83	7.67	19.2%	14.2%	2.5%
10 Baseline specification, without the $570 price group	0.641	460.57	0.740	202.09	9.77	17.0%	14.5%	1.8%

Panel B: Parameter estimates from different specifications

	Demand Equation				Average Cost Equation			
	alpha		beta		gamma		delta	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
1 Baseline (family coverage, no state fixed-effects)	0.940	(0.123)	-0.00070	(0.00032)	391.7	(26.8)	0.155	(0.064)
Robustness to demand estimates								
2 Probit demand	1.149	(0.316)	-0.00183	(0.00080)	------ same as Baseline ------			
3 Linear demand, constrained to go through (Q,P)=(1,$0)	1.000	(imposed)	-0.00084	(0.00005)	------ same as Baseline ------			
4 Linear demand, constrained to go through (Q,P)=(0,$800)	1.333	(imposed)	-0.00167	(0.00005)	------ same as Baseline ------			
5 Quadratic demand, constrained to go through (1,$0) and (0,$800)[f]	1.000	(imposed)	{ -0.00039	(imposed)	------ same as Baseline ------			
			-1.1E-06	(1.2E-07)				
Robustness to tax subsidy								
6 Baseline specification, but accouting for pre-tax premiums	0.940	(0.123)	-0.00107	(0.00048)	391.7	(26.8)	0.239	(0.098)
Robustness to sample and source of variation								
7 State fixed-effects included (in both regressions)	0.919	(0.167)	-0.00065	(0.00040)	414.8	(37.0)	0.099	(0.090)
8 State fixed-effects and demographics included (in both regressions)	0.917	(0.170)	-0.00063	(0.00040)	394.5	(36.9)	0.104	(0.091)
9 All coverage tiers, no state fixed-effects[e]	0.848	(0.109)	-0.00059	(0.00032)	374.8	(22.8)	0.137	(0.062)
10 Baseline specification, without the $570 price group	0.818	(0.195)	-0.00038	(0.00050)	406.3	(34.0)	0.118	(0.081)

Table reports results from alternative specifications. Panel B reports parameter estimates, and Panel A reports the (corresponding) implications for welfare analysis. Row 1 replicates the results from the baseline specification (as in Table III), rows 2-5 report specifications that change the functional form of demand. Row 6 re-estimates the baseline specification with the price in both the demand and cost equation multiplied by 0.65 (one minus the average marginal tax rate in the sample). Row 7 includes state fixed effects in both the demand and cost equations, and row 8 also controls for employee characteristics (listed in Table I). Row 9 increases the sample to include employees in all four coverage tiers. Row 10 tries to assess sensitivity to dropping the $570 price group, which is the greatest outlier (see Figure V). Standard errors (in parentheses) allow for an arbitrary variance-covariance matrix within each state.

[a] Graphically, this is the area of triangle CDE (see Figure I).

[b] This is triangle CDE divided by $0.3Q_{eff}(P_{eq}-P_{eff})$.

[c] Graphically, this is the area of triangle CDE divided by the area of triangle EGH (see Figure I).

[d] Graphically, this is the area of triangle CDE divided by the area of triangle ABE (see Figure I).

[e] N=7,263 for demand analysis, 4,622 for cost analysis; mean dependent variables are 0.64 (D) and \$424 (c), respectively. We include (de-meaned) indicator variables for the coverage tier in both the demand and cost equations (not shown); we multiply p and c by two for employees in the "employee only" coverage tier.

[f] In the quadratic demand specification, the top reported coefficient of beta is the coefficient on the linear term, while the second is the coefficient on the quadratic term.

Table A3: Potential sample selection

Dependent variable:	1 if "outside good" was chosen, 0 otherwise		
	"Outside Good" does not include "opt out"		"Outside good" does include "opt out"
	Family coverage tier only (1)	All coverage tiers (2)	All coverage tiers (3)
Relative price	-0.0000093	-0.000021	0.000002
	(0.00035)	(0.00040)	(0.000003)
	[0.98]	[0.96]	[0.66]
Constant	0.287	0.292	0.296
	(0.1580)	(0.1150)	(0.1580)
	[0.08]	[0.02]	[0.07]
Mean dependent variable	0.283	0.300	0.359
Number of obs.	5,271	10,386	11,325

The table reports results of estimating a variant of the demand equation shown in equation (11). The dependent variable is an indicator variable that takes the value of 1 if the employee chose any of the "outside options" and 0 if the employee chose either contract H or contract L. The "relative price" variable is, as in Table II, the relative price of contract H compared to contract L. In columns (1) and (2) the "outside good" includes two lower coverage PPOs, a Health Reimbursement Account PPO, and an HMO. The sample in column (1) is limited to family coverage. The sample in column (2) includes all coverage tiers. We therefore include (de-meaned) indicator variables for the coverage tier (not shown) and multiply the price variable by two for employees in the "employee only" coverage tier. In column (3) the "outside good" definition is expanded to also include employees who opt out of coverage. Since coverage tier is not known for these employees, we include all employees regardless of coverage tier and do not include indicator variables for coverage tier. We define the price variable as the relative price associated with family coverage (regardless of the actual tier chosen, if known). All estimates are generated by OLS, standard errors (in parentheses) allow for an arbitrary variance-covariance matrix within each state, ; p-values are in [square brackets].