Externalities and Taxation of Supplemental Insurance:
A Study of Medicare and Medigap
Online Appendix

Marika Cabral Neale Mahoney

May 7, 2018

A Medigap Plans: Plan Features and Enrollees by Plan Letter

The form and pricing of Medigap policies are regulated by the federal government. During our sample period, firms were permitted to sell standardized policies labeled A-J. Table A1 describes the features of these different Medigap policies. As one can see, all the policies contain the "basic benefits," which include coverage of Part A copays and deductibles, Part B coinsurance, blood, and additional lifetime hospital days. Much of the differentiation among the plans is for niche services such as home health care and foreign travel emergencies.

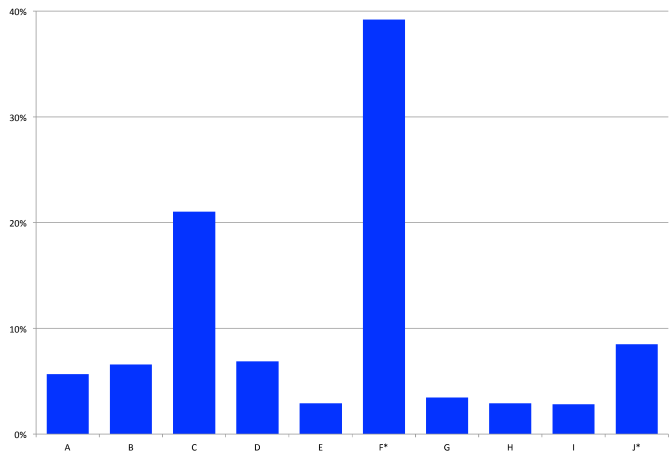
Table A1: Medigap Benefits by Plan Letter

	Medigap Plan Letter									
	A	B	C	D	E	F*	G	H	I	J*
Basic Benefits	X	X	X	X	X	X	X	X	X	X
Part A Copays and Deductible										
Part B Coinsurance										
Blood										
Additional Lifetime Hospital Days										
SNF Coinsurance		X	X	X	X	X	X	X	X	X
Part B Deductible		X				X				X
Part B Excess Charges						X	80%			
Foreign Travel Emergency		X	X	X	X	X	X	X	X	X
Home Health Care				X			X		X	X
Prescription Drugs								X	X	X
Preventive Medical Care				X						X

Notes: Table shows Medigap plan benefits by plan letter. The "basic benefits" are provided by all plans. According to federal regulations, firms that participate in the Medigap market must offer Plan A and either Plan C or Plan F.
*Plans F and J have high-deductible options that require beneficiaries to pay $1,580 before receiving Medigap benefits that year. These plans are rarely offered and have very few enrollees.

Figure A1 illustrates the distribution of Medigap enrollees by plan letter. This distribution is calculated from self-reported Medigap plan letter information from the MCBS (which is reported by roughly half of the respondents who report having Medigap coverage). As one can see, Plan C and Plan F are the most popular plans. Federal government regulations required firms that offered any Medigap policy to offer two options as a subset of the available plans: Plan A and either Plan C or Plan F.

Figure A1: Medigap Enrollment by Plan Letter



Notes: Figure displays enrollment by plan letter. This histogram is constructed using data from the 1992-2005 Medicare Current Beneficiary Survey (MCBS). A Medigap plan letter is reported by approximately half of the MCBS respondents who report having a Medigap policy.

B Supplemental Insurance in MCBS and NHIS datasets

To investigate the elasticity of Medigap enrollment, we use data from two surveys: the Medicare Current Beneficiary Survey (MCBS) and the National Health Interview Survey (NHIS). Below, we describe how we translate the variables in these surveys into the insurance dependent variables we use in the demand estimation: Medigap (from MCBS) and supplemental insurance (from MCBS and NHIS).

MCBS. MCBS insurance variables are available in the "ric 4" data file for each year. We code individuals as having Medigap if they report having private coverage and report the plan is "self-purchased" and either purchased directly or through AARP. We code individuals as having supplemental coverage if we can infer that they have any source of supplemental coverage, including Medicaid, Medicare Advantage, Medigap, or RSI. Specifically, the following MCBS variables are used in coding individual insurance status: d_phi, d_hmo, d_caid, d_obtnp1-5. [1]

NHIS. NHIS insurance variables are available in the "personx" data file for each year. Relative to the MCBS, the NHIS has fewer survey questions regarding sources of coverage, and the NHIS survey responses are not verified against administrative data. We code individuals as having supplemental insurance if we can infer that they have any source of supplemental coverage, including Medicaid, Medicare Advantage, Medigap, or RSI. Specifically, the following NHIS variables are used in coding individual insurance status: mchmo, medicare, plnpay21, plnpay22, private, plnwrkn1, plnwrk2, medicaid.

C Robustness of Demand Results

Robustness to Alternative Control Variables In the following table, we display our demand results with alternative sets of controls. The table shows the estimates from the baseline specifica-

[1] The characterization leads to roughly the same market shares as displayed in GAO (2001).

tion for reference (as in Table 5). All specifications include year fixed effects, local medical market fixed effects, and controls for Medicare geographic payment adjustments. The "Fewer Controls" specification includes no additional controls, and the "Baseline Controls" specification includes demographic controls for sex, race, and age. The "More Controls" specification includes demographic controls as well as controls for the incidence of chronic conditions including arthritis, heart disease, diabetes, non-skin cancer, and previous heart attack. The table displays the results for the dependent variables indicating Medigap (in the MCBS) and any supplemental coverage (in the NHIS and the MCBS). Regardless of which set of controls are used, the results are qualitatively and quantitatively very similar. The results indicate that Medigap enrollment is price-sensitive, and the implied elasticity from the combined specification is in the range of -1.5 to -1.8. Within the MCBS, the effects on Medigap and any supplemental insurance are very similar, consistent with the evidence from the administrative data on the lack of substitution into alternative coverage based on our variation.

Table C1: Demand: Robustness to Alternative Controls

	Baseline Controls		Fewer Controls		More Controls		Mean of
	Est	Std. Err.	Est	Std. Err.	Est	Std. Err.	Dep Var
All Beneficiaries							
Combined MCBS+NHIS							
Supplemental Coverage (HSA level)	-0.048	(0.023)	-0.046	(0.023)	-0.048	(0.024)	0.85
Supplemental Coverage (HRR level)	-0.040	(0.016)	-0.036	(0.015)	-0.042	(0.016)	0.85
MCBS Alone							
Supplemental Coverage (HSA level)	-0.066	(0.038)	-0.068	(0.038)	-0.064	(0.040)	0.90
Supplemental Coverage (HRR level)	-0.071	(0.028)	-0.068	(0.026)	-0.073	(0.028)	0.90
Medigap (HSA level)	-0.083	(0.060)	-0.080	(0.064)	-0.079	(0.060)	0.36
Medigap (HRR level)	-0.090	(0.049)	-0.088	(0.047)	-0.092	(0.048)	0.36
NHIS Alone							
Supplemental Coverage (HSA level)	-0.031	(0.027)	-0.026	(0.025)	-0.032	(0.027)	0.79
Supplemental Coverage (HRR level)	-0.010	(0.016)	-0.006	(0.016)	-0.012	(0.016)	0.79
Controls							
Year and Local Medical Market Fixed Effects	X		X		X		
Demographic	X				X		
Chronic Conditions					X		

Notes: Table shows estimates from regressions of insurance coverage indicators on leave-out costs, HSA or HRR fixed effects, and controls as indicated in the table above (see Section II, Equation 3). The analysis uses the MCBS and NHIS data from 1992 to 2005, using a sample definition analogous to Panel A of Table 2. The dependent variable in the Supplemental Coverage specifications is an indicator for Medigap, Medicare Advantage, Medicaid, or RSI coverage. The HRR-level first stage ranges from 0.24 to 0.25 across specifications (Appendix Table C2) and we scale the HRR demand estimates by 4 to make them comparable to the HSA-level estimates, which have first-stage of 0.94 to 1.1 across specifications. Standard errors are clustered at the HSA or HRR level depending on the specification. Dollar values are inflation-adjusted to 2005 using the CPI-U.

First-stage at HRR level. The following table presents the first stage at the HRR level. The estimates show that HRR-level leave-out costs are predictive of premiums, with a coefficient around 0.25. Recall that the analogous coefficient at the HSA level was approximately 1. The reason the HRR-level coefficient is smaller is that HRRs are substantially larger than HSAs and therefore the geographic areas used to calculate HRR-level leave-out costs are substantially smaller than the areas used to calculate leave-out costs at the HSA level. Because of these smaller areas, HRR-level leave-out costs are more noisy predictors of state-level costs and thus more noisy predictors of premiums, attenuating the coefficient on HRR-level leave-out costs towards zero. Because of the attenuated coefficient, the demand results at the HRR level need to be scaled up by a factor of four to be comparable with the HSA-level estimates. This is what is done in reporting the results for the demand coefficients.

Table C2: Premiums: Regressions of Annual Medigap Premiums on Leave-Out Costs at HRR level

	Dep. Var.: Annual Medigap Premiums		
	Plans A-J	Plan C	Plan F
	(1)	(2)	(3)
Leave-Out Costs	0.241	0.244	0.249
	(0.057)	(0.059)	(0.068)
HRR FE	X	X	X
Insurer FE	X	X	X
Plan FE	X		
R-Squared	0.917	0.805	0.838
Mean of Dep Var	1,538	1,431	1,477
N	44,765	6,246	6,397

Notes: Table shows estimates from regressions of annual Medigap premiums on the leave-out costs instrument, HRR fixed effects, plan fixed effects, insurer fixed effects, and controls for GAF/OWI adjustment factors. The first column displays results from a specification that includes all plans offered by United Healthcare and Mutual of Omaha, the two largest insurers. The second and third columns restrict attention to the most popular plans offered by these companies, Plan C and Plan F, respectively. Observations are at the HRR-state-plan-company level. Standard errors are clustered at the HRR level. Dollar values are inflation-adjusted to 2005 using the CPI-U.

OLS As we discussed in Section II of the paper, we only have premium data for plans offered in the year 2000, and we do not have data on plan market share. Thus, it is not possible to estimate the OLS analogue of the IV using a market-share weighted mean premium measure. However, below we report OLS regression estimates using three feasible premium measures constructed from the year 2000 premium data for the two largest national Medigap insurers (United Healthcare and Mutual of Omaha): mean unweighted premium across all plans, mean premium for Plan C, and mean premium for Plan F. Nationwide, Plans C and F are the most commonly purchased Medigap plans (see Figure A1).

Table C3: Demand: OLS with Various Premium Measures

	Premium (Hundreds)			Mean of
	Est	Std. Err.	P-Value	Dep Var
Unweighted Mean Premium, All Plans				
Supplemental Coverage	-0.009	(0.009)	0.31	0.90
Unweighted Mean Premium, Plan C				
Supplemental Coverage	-0.040	(0.017)	0.02	0.90
Unweighted Mean Premium, Plan F				
Supplemental Coverage	-0.016	(0.008)	0.05	0.90

Notes: Table shows estimates from regressions of insurance coverage indicators on the premium measure noted in the table, HSA fixed effects, and controls for age, sex, and GAF/OWI adjustment factors (see Section II, Equation 3). The analysis uses the MCBS data from 1992 to 2005, using a sample definition analogous to Panel A of Table 2. The dependent variable indicates if an individual has any supplemental coverage: Medigap, Medicare Advantage, Medicaid, or RSI coverage. Standard errors are clustered at the HSA level. Dollar values are inflation-adjusted to 2005 using the CPI-U.

D Control Variables in CMS Data

We include controls for individual chronic conditions in several specifications to improve the precision of our estimates. The chronic condition information comes from the CMS Beneficiary Summary File. The chronic condition controls we include are dummy variables that indicate when the following conditions are present:

- Acute Myocardial Infarction (AMI)
- Alzheimer's Disease (ALZH)
- Alzheimer's Disease and Rltd Disorders or Senile Dementia (ALZHDMTA)
- Atrial Fibrillation (ATRIALFB)
- Cataract (CATARACT)
- Chronic Kidney Disease (CHRNKIDN)
- Chronic Obstructive Pulmonary Disease (COPD)
- Heart Failure (CHF)
- Diabetes (DIABETES)
- Glaucoma (GLAUCOMA)
- Hip/Pelvic Fracture (HIPFRAC)
- Ischemic Heart Disease (ISCHMCHT)
- Depression (DEPRESSN)
- Osteoporosis (OSTEOPRS)
- Rheumatoid Arthritis or Osteoarthritis (RA_OA)
- Stroke or Transient Ischemic Attack (STRKETIA)
- Breast Cancer (CNCRBRST)
- Colorectal Cancer (CNCRCLRC)
- Prostate Cancer (CNCRPRST)
- Lung Cancer (CNCRLUNG)
- Endometrial Cancer (CNCRENDM)
- Anemia (ANEMIA)
- Asthma (ASTHMA)

- Hyperlipidemia (HYPERL)
- Benign Prostatic Hyperplasia (HYPERP)
- Hypertension (HYPERT)
- Acquired Hypothyroidism (HYPOTH)

The CMS corresponding variable used to derive each of these indicator variables is included in the list above in parentheses after each chronic condition. For more information on the CMS algorithm for determining whether these conditions are present, see the documentation at: http://www.resdac.org/cms-data/files/mbsf/data-documentation.

E Alternative Specifications

The baseline specifications reported in the text include controls for age, sex, race, chronic conditions, and log GAF/OWI adjustment factors. In Table E1, we report the results for the utilization dependent variables when chronic health condition controls are omitted. Overall, the results are qualitatively similar as when the chronic health condition controls are included.

Table E1: Utilization: Regressions of Annual Medicare Utilization on Leave-Out Costs, Without Health Controls

	Leave-Out Costs (Hundreds)			Mean of	Implied Medigap Effect	
Dependent Variable	Est	Std. Err.	P-Value	Dep. Var.	Level	%
Part B Events	-0.3210	(0.1990)	0.106	25.81	6.69	25.9%
Imaging Events	-0.0561	(0.0337)	0.096	3.99	1.17	29.3%
Testing Events	-0.3400	(0.1710)	0.047	11.41	7.08	62.1%
Total RVUs	-0.9550	(0.4970)	0.055	70.77	19.90	28.1%
Part A Days	-0.0354	(0.0218)	0.105	2.10	0.74	35.1%
Part A Stays	0.0002	(0.0025)	0.931	0.34	0.00	-1.3%
SNF Days	0.0246	(0.0251)	0.327	1.37	-0.51	-37.3%
SNF Stays	0.0009	(0.0011)	0.414	0.06	-0.02	-29.9%

Notes: Table displays estimates from regressions of annual Medicare utilization on leave-out costs, HSA fixed effects, and controls for age, sex, race, and GAF/OWI adjustment factors (see Section II Equation 4). Each row displays the results from a separate regression. The implied Medigap effect is calculated by dividing the estimate by the coefficient on leave-out costs from the baseline demand specification. This analysis draws on data from the pooled 1999-2005 CMS Beneficiary Summary File, CMS Denominator File, and CMS Carrier File (for RVU analysis). This analysis uses the baseline sample described in Panel B of Table 2 (N=23,708,295 for the RVU measure; N=130,895,953 for all other measures). The difference between these results and those presented in Table 6 is that this specification excludes health controls. Standard errors are clustered at the HSA level. Dollar values are inflation-adjusted to 2005 using the CPI-U.

Table E2 reports the results when the baseline specification for the payment dependent variables is estimated omitting the chronic health condition controls. The results are less statistically precise when these controls are omitted. However, these results are statistically indistinguishable from the point estimates in the baseline specification. It is perhaps not surprising that the health controls are important for precision as the R-squared increases from 0.03 without health controls to 0.43 with health controls for the "Medicare Payments" specification.

Table E2: Payments: Regressions of Annual Medicare Payments on Leave-Out Costs, Without Health Controls

	Leave-Out Costs (Hundreds)			Mean of	Implied Medigap Effect	
Dependent Variable	Est	Std. Err.	P-Value	Dep. Var.	Level	%
Medicare Payments	-9.98	(33.45)	0.766	6,291	207.83	3.3%
Part A Payments	-4.70	(22.18)	0.832	3,021	97.87	3.2%
Part B Payments	-11.36	(16.76)	0.498	2,648	236.73	8.9%
SNF Payments	6.90	(6.42)	0.283	399	-143.84	-36.0%

Notes: Table displays estimates from regressions of annual Medicare payments on leave-out costs, HSA fixed effects, and controls for age, sex, race, and GAF/OWI adjustment factors (see Section II, Equation 4). Each row displays the results from a separate regression. The implied Medigap effect is calculated by dividing the estimate by the coefficient on leave-out costs from the baseline demand specification. This analysis draws on data from the pooled 1999-2005 CMS Beneficiary Summary File and CMS Denominator File. This analysis uses the baseline sample described in Panel B of Table 2 (N=130,895,953). All dependent variables are top-coded at $64,000. The difference between these results and those presented in Table 7 is that this specification excludes health controls. Standard errors are clustered at the HSA level. Dollar values are inflation-adjusted to 2005 using the CPI-U.

F Heterogeneity

The baseline analysis presented in the text uses administrative cost and utilization data from 1999-2005. Below we present our baseline Medicare spending regression estimated year-by-year. While the subsample estimates are a bit more noisy, overall the year-by-year estimates line up with the estimates on the entire sample.

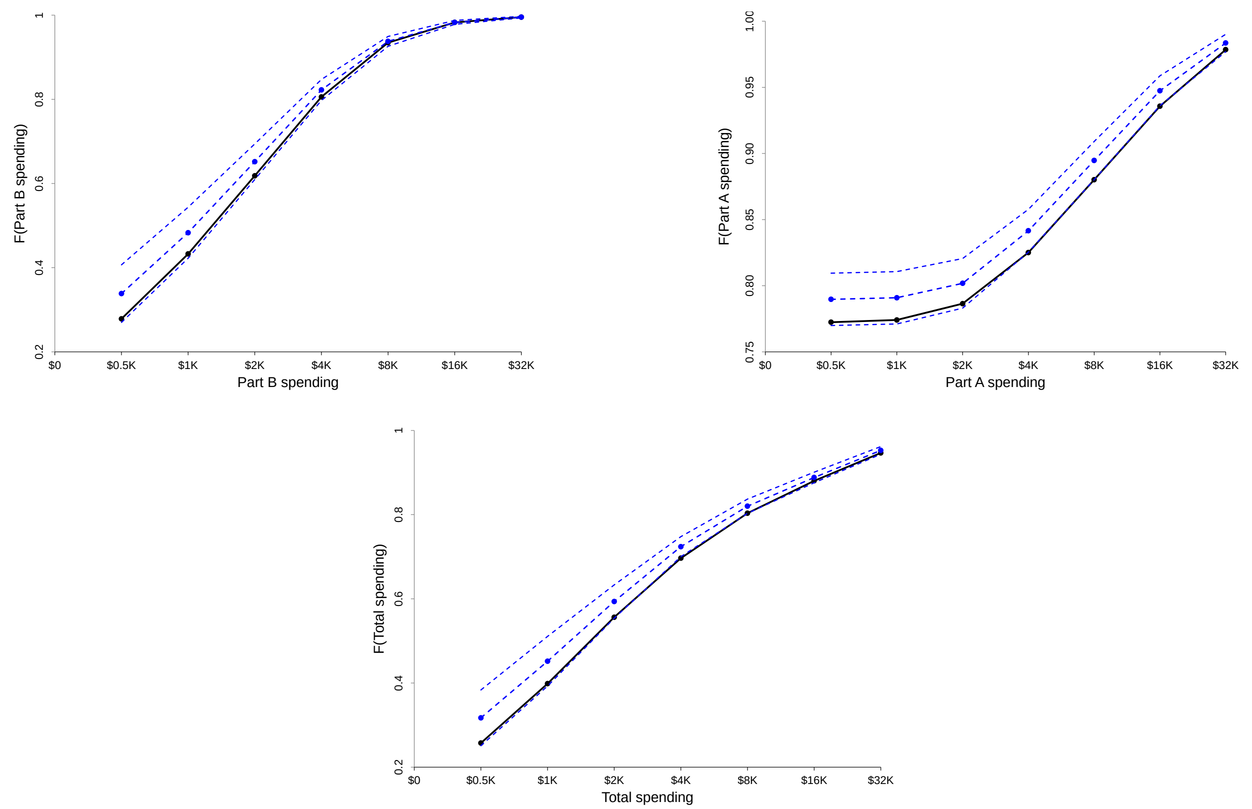
Table F1: Annual Medicare Payment Regressions by Year

	Leave-Out Costs (Hundreds)			Mean of	Implied Medigap Effect		
Dependent Variable	Est	Std. Err.	P-Value	Dep. Var.	Level	%	N
Annual Medicare Payments							
Baseline (1999-2005)	-67.0	(33.1)	0.043	6,291	1,396	22.2%	130,895,953
By year							
1999	-30.4	(42.0)	0.469	6,049	633	10.5%	17,896,807
2000	-80.2	(39.4)	0.042	5,975	1,671	28.0%	17,990,486
2001	-69.6	(39.7)	0.080	6,195	1,450	23.4%	18,464,210
2002	-58.1	(41.6)	0.163	6,305	1,210	19.2%	18,968,995
2003	-20.3	(35.9)	0.572	6,398	423	6.6%	19,196,098
2004	-48.4	(41.5)	0.243	6,556	1,008	15.4%	19,319,846
2005	-83.8	(40.0)	0.036	6,518	1,746	26.8%	19,059,511

Notes: Table displays estimates from regressions of Medicare payments on leave-out costs, HSA fixed effects, and controls for age, sex, race, health risk, and GAF/OWI adjustment factors (see Section II, Equation 4). Each row displays the results from a separate regression. The baseline estimates using pooled data from 1999-2005 are displayed in the first row, while the subsequent rows present estimates from the same specification estimated separately by year. The implied Medigap effect is calculated by dividing the estimate by the coefficient on leave-out costs from the baseline demand specification. This analysis draws on data from the pooled 1999-2005 CMS Beneficiary Summary File and CMS Denominator File. This analysis uses the baseline sample described in Panel B of Table 2 (N=130,895,953 for all years). All dependent variables are top-coded at $64,000. Standard errors are clustered at the HSA level. Dollar values are inflation-adjusted to 2005 using the CPI-U.

In the text, we focus on the mean effect of Medigap premiums on Medicare payments, as this is the relevant object for evaluating the effect of a tax on Medigap. Below we present additional graphical evidence of the effect of Medigap premiums on the distribution of Medicare payments. Figure F1 shows the effect of a $1,000 increase in leave-out costs on the CDFs of Part A, Part B, and total Medicare payments. Solid lines show the CDF of payments in each category.[2] Dashed lines depict the effect of a $1,000 increase in leave-out costs. The lines are calculated using the coefficient on leave-out costs from regressions of the form $\Pr(\text{Payments}_{ijk} < X) = \gamma_c \text{Leave-out costs}_{jk} + \gamma_k + X'_{ijk}\gamma_X + \mu_{ijk}$ where $X = 500, 1{,}000, \ldots 32{,}000$. Dotted lines show the 95% confidence intervals of these estimates, calculated using standard errors clustered at the HSA-level.

Figure F1: Effect on CDF of Annual Medicare Payments



Notes: Figure shows the impact of a $1,000 increase in leave-out costs on the CDF of annual Part B payments, annual Part A payments, and total annual Medicare payments. The solid lines show the empirical CDF of payments. The dashed lines show the estimated CDF under an $1,000 increase in leave-out costs. These CDFs are constructed using the coefficient on leave-out costs from regressions of the form $\Pr(\text{Payments} < X) = \gamma_c \text{Leave-out costs}_{jk} + \gamma_k + X'_{ijk}\gamma_X + \mu_{ijk}$ for $X = 500, 1{,}000, \ldots 32{,}000$. Dotted lines show the 95% confidence intervals of these estimates. This analysis draws on data from the pooled 1999-2005 CMS Beneficiary Summary File and CMS Denominator File. This analysis uses the baseline sample described in Panel B of Table 2 (N=130,895,953 for all years). Standard errors are clustered at the HSA level. Dollar values are inflation-adjusted to 2005 using the CPI-U.

G Premium Variation and Mortality

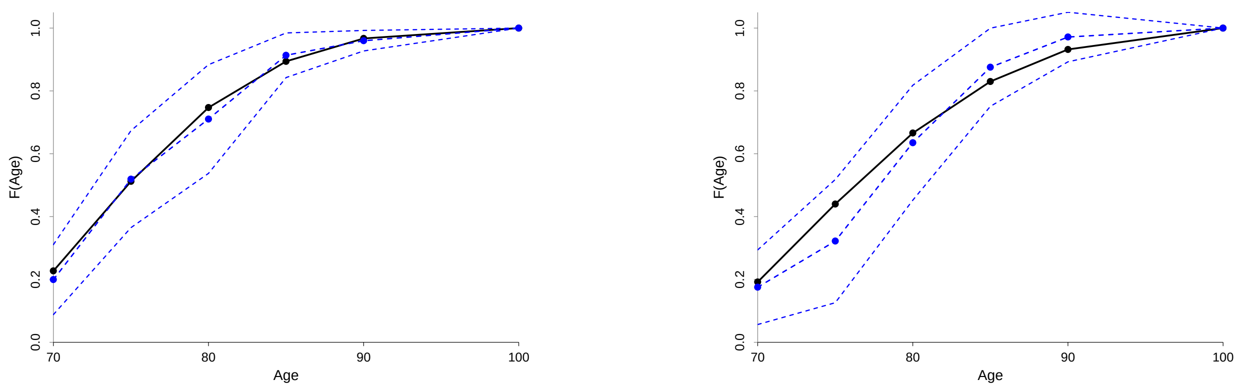
We can analyze the effect of Medigap on mortality by examining the effect of the instrument on the cross-sectional age distribution of Medicare beneficiaries. If Medigap reduces mortality,

[2] The distribution is censored at $32,000 per year.

then higher leave-out costs, and the corresponding lower Medigap take-up, should lead to earlier death, shifting the age distribution in an inward direction.

Figure G1 displays the impact of a \$10 increase in leave-out costs on the cross-sectional age distribution. Solid lines show the empirical CDF of age. Dashed lines show the estimated CDF under a \$10 increase in leave-out costs. These CDFs are constructed using the coefficient on leave-out costs from regressions of the form $\text{Fraction}(\text{Age} < X)_{zjk} = \gamma_c \text{Leave-out costs}_{jk} + \gamma_k + \mu_{zjk}$ for $X = 70, 75, \ldots 100$, where zipcode-year observations are weighted by the population residing within the zipcode. Dotted lines show the 95% confidence intervals of these estimates. Overall, the plots show that the instrument has no detectable effect on the age distribution of Medicare beneficiaries. Although this evidence is consistent with Medigap having no mortality effect, our research design does not have the power to detect small to moderate effects on mortality.

Figure G1: Effect on Cross-Sectional Age Distribution



(a) Men

(b) Women

Notes: Figure shows the impact of a \$10 increase in leave-out costs on the cross-sectional age distribution. Solid lines show the empirical CDF of age. Dashed lines show the estimated CDF under a \$10 increase in leave-out costs. These CDFs are constructed using the coefficient on leave-out costs from zipcode-year level regressions of the form $\text{Fraction}(\text{Age} < X)_{zjk} = \gamma_c \text{Leave-out costs}_{jk} + \gamma_k + \mu_{zjk}$ for $X = 70, 75, \ldots 100$, where zipcode-year observations are weighted by the population residing within the zipcode. Dotted lines show the 95% confidence intervals of these estimates. This analysis draws on data from the pooled 1999-2005 CMS Denominator File. This analysis uses the baseline sample described in Panel B of Table 2 (N=130,895,953 for all years). Standard errors are clustered at the HSA level. Dollar values are inflation-adjusted to 2005 using the CPI-U.

H Robustness to Level of Clustering

Table H1 below displays our main utilization and payment regressions along with standard errors utilizing various levels of clustering: HSA and State, HSA, State, HSA-State, 5-digit Zipcode, and individual. For each specification, the table notes the cluster-adjusted standard error and p-value, along with the number of clusters. As expected, the precision of our estimates goes up as we cluster on finer levels.

It is important to note that there is a trade-off between clustering at different levels: higher levels of clustering allow for more flexible correlation among observations, while lower levels of clustering may be more reliable in finite sample with finitely many clusters. We use HSA-level clustering in the baseline specifications reported in the main text (also reported in columns 2 and 3 of Table H1). There are a few reasons for this choice. First, HSA is arguably the most important level of clustering, because individuals within an HSA see a common set of medical providers and there is a well-established literature documenting the importance of medical providers as a determinant of medical spending (e.g., Finkelstein, Gentzkow and Williams (2016), Wennberg,

Table H1: Medicare Payments and Utilization Regressions: Robustness to Alternative Level of Clustering

	Leave-Out Costs (Hundreds)												
	Level of Clustering for Standard Errors												
		Baseline Cluster Level		Alternative Cluster Level									
		HSA		Individual		5-digit Zipcode		HSA-State		State		Multiway: HSA, State	
Dependent Variable	Est	Std. Err.	P-Value	Std. Err.	P-Value	Std. Err.	P-Value	Std. Err.	P-Value	Std. Err.	P-Value	Std. Err.	P-Value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Utilization													
Part B Events	-0.418	(0.181)	0.021	(0.029)	<0.001	(0.107)	<0.001	(0.154)	0.006	(0.157)	0.011	(0.235)	0.082
Imaging Events	-0.081	(0.032)	0.012	(0.005)	<0.001	(0.018)	<0.001	(0.027)	0.003	(0.027)	0.005	(0.048)	0.100
Testing Events	-0.409	(0.147)	0.005	(0.019)	<0.001	(0.099)	<0.001	(0.125)	0.001	(0.135)	0.004	(0.177)	0.026
Total RVUs	-1.290	(0.496)	0.009	(0.188)	<0.001	(0.289)	<0.001	(0.410)	0.002	(0.342)	<0.001	(0.629)	0.046
Part A Days	-0.062	(0.019)	0.001	(0.006)	<0.001	(0.011)	<0.001	(0.016)	<0.001	(0.017)	0.001	(0.024)	0.015
Part A Stays	-0.00395	(0.00214)	0.065	(0.00072)	<0.001	(0.00129)	0.002	(0.00182)	0.030	(0.00168)	0.023	(0.00270)	0.150
SNF Days	0.012	(0.020)	0.552	(0.008)	0.111	(0.014)	0.406	(0.017)	0.492	(0.024)	0.622	(0.032)	0.708
SNF Stays	0.00026	(0.00085)	0.761	(0.00030)	0.396	(0.00055)	0.640	(0.00070)	0.713	(0.00083)	0.759	(0.00120)	0.831
Payment													
Medicare Payments	-67.0	(33.1)	0.043	(10.4)	<0.001	(19.1)	<0.001	(27.8)	0.016	(26.1)	0.014	(42.0)	0.117
Part A Payments	-47.6	(22.8)	0.037	(7.5)	<0.001	(13.5)	<0.001	(19.6)	0.015	(16.6)	0.006	(28.9)	0.107
Part B Payments	-21.8	(15.5)	0.159	(4.9)	<0.001	(8.2)	0.007	(12.4)	0.079	(12.4)	0.085	(17.1)	0.209
SNF Payments	3.44	(5.3)	0.513	(2.0)	0.083	(3.6)	0.343	(4.6)	0.455	(6.3)	0.585	(8.5)	0.688
Number of Clusters		3,121		21,989,766		31,003		3,352		45		3,121 and 45	

Notes: Table displays estimates from regressions of annual Medicare utilization and annual Medicare payments on leave-out costs, HSA fixed effects, and controls for age, sex, race, health risk, and GAF/OWI adjustment factors (see Section II Equation 4). Each row in the table above displays estimates for a different dependent variable. Columns present standard errors (and associated p-values) with alternative levels of clustering. This analysis draws on data from the pooled 1999-2005 CMS Beneficiary Summary File, CMS Denominator File, and CMS Carrier File (for RVU analysis). This analysis uses the baseline sample described in Panel B of Table 2 (N=23,708,295 for the RVU measure; N=130,895,953 for all other measures). Dollar values are inflation-adjusted to 2005 using the CPI-U.

Fisher and Skinner (2002)). Second, the HSA-level nests the level of the variation of the leave-out costs instrument, as our instrument is constructed at the HSA-State level. Third, based on our analysis in the table above, clustering at the HSA level is generally more conservative than clustering at the state level.

I Definition of Urgent Procedures

I.A Betos Code Characterization from Clemens and Gottlieb (2013)

We follow the categorization used by Clemens and Gottlieb (2014) to group Part B RVUs by BETOS code to determine which procedures are for less discretionary care. Specifically, we define urgent procedures as Part B claims associated with the following BETOS codes:

- P1A: Major procedure—breast
- P1B: Major procedure—colectomy
- P1C: Major procedure—cholecystectomy
- P1D: Major procedure—turp
- P1F: Major procedure—hysterectomy
- P1G: Major procedure—other
- P2B: Major procedure, cardiovascular—aneurysm repair
- P3A: Major procedure, orthopedic—hip fracture repair
- P4A: Eye procedure—corneal transplant

- P4C: Eye procedure—retinal detachment
- P5C: Ambulatory procedure—groin hernia repair
- P7A: Oncology—radiation therapy
- P7B: Oncology—other
- P9A: Dialysis services

I.B Weekend Versus Weekday Daily Frequency Characterization from Card, Dobkin and Maestas (2009)

Card, Dobkin and Maestas (2009) characterize urgent hospitalizations by inspecting the weekend versus weekday daily frequency of ICD-9 codes for hospital admissions originating in the ER. We consider two definitions urgent hospitalizations based on this characterization. For the first definition, we define a procedure as urgent if it is listed in Table I of Card, Dobkin and Maestas (2009) as one of the ten highest frequency urgent ICD-9 diagnoses based on their data and characterization. Below is the list of procedures that this first definition encompasses.

- Obstructive chronic bronchitis with acute exacerbation
- Respiratory failure
- AMI of other inferior wall (1st episode)
- AMI of other anterior wall (1st episode)
- Intracerebral hemorrhage
- Chronic airway obstruction, n.e.c.
- Fracture of neck of femur intertrochanteric section
- Cerebral artery occlusion, unspecified
- Convulsions unknown cause
- Asthma, unspecified with status asthmaticus

For the second definition, we apply the same procedure as Card, Dobkin and Maestas (2009) to the 2002 CMS MedPAR data to identify urgent procedures. Specifically, we construct the fraction of hospitalizations originating from the ER during the weekend for each ICD-9 code. We then define a hospitalization as urgent if the T-stat on this fraction being equal to $\frac{2}{7}$ is less than or equal 0.3713 (the 10th percentile of the distribution of T-stats).[3] Below are the descriptions of the ten highest frequency ICD-9 codes that are characterized as urgent through this second methodology:

- Escherichia coli infections
- Paralytic ileus
- Home accidents (Accident in home)
- Acute pancreatitis
- Other abnormal blood chemistry (Abn blood chemistry NEC)
- Diverticulitis of colon (without mention of hemorrhage) (Dvrtcli colon w/o hmrhg)
- Infection with microorganisms resistant to penicillins (Inf mcrg rstn pncllins)
- Benign neoplasm of colon (Benign neoplasm lg bowel)
- Other closed transcervical fracture of neck of femur (Fx femur intrcaps NEC-cl)
- Acute myocardial infarction of other inferior wall

[3]Note that this definition of urgent procedures is more conservative than that in Card, Dobkin and Maestas (2009). Card, Dobkin and Maestas (2009) define a procedure as urgent if the T-stat is less than 0.965.

J Robustness to Spatial Trends in Utilization

Many determinants of health care utilization vary continuously over geography, including provider choice, environmental factors, and behavioral factors. If these determinants of health care utilization are correlated with the instrument, our identification assumption will not hold. We address this concern in three ways. First, we re-estimate the baseline specification, restricting the sample of individuals within cross-border HSAs to be those within a very short distance of the state boundary. The idea behind this sample restriction is that if there are spatial trends in health care utilization (driven by characteristics such as provider choice and demographics), then those individuals who live closest to one another are the best controls for one another. Table J1 reports the results. The point estimates remain statistically significant and similar in magnitude when we concentrate on the sample within 30 kilometers of state boundaries.[4,5] This is reassuring as this restricted sample contains individuals who are most similar to one another in terms of continuously trending unobservables.

Second, we verify that our estimates are robust to spatially trending omitted variables by estimating a specification with carefully defined placebo borders. Specifically, we partition each HSA-state segment in cross-border HSAs into two areas: the *border area* within 20 km of the state boundary and the *near border area* consisting of the remainder of the HSA-state. The placebo border is then the division between these two areas, meaning that placebo border is entirely internal to the state in question. We then assign the *border area* a counterfactual instrument equal to the instrument of the neighboring state, while the *near border area* has the true value of the instrument as in our baseline estimation. With this newly defined instrument determined by the placebo border, we then run the same regressions as in the baseline specification replacing the HSA fixed effects with HSA-state fixed effects. The results are reported in Table J1. If the baseline results are not picking up the causal effect of Medigap but instead reflecting unrelated spatial trends in medical spending, then one would expect the coefficient from this specification to be the same as in our baseline specification. In contrast to the significant results in our baseline estimation, we see that the coefficient in this specification is statistically indistinguishable from zero (with a p-value of 0.54). This test reveals that our estimated effect of Medigap is not simply reflecting unrelated, continuous spatial trends in medical utilization.

Third, we evaluate the robustness of our estimates to an alternative leave-out costs measure that drops individuals in a "buffer zone" around the cross-border HSAs. The Dartmouth Atlas aggregates HSAs into larger HRRs based on markets for tertiary medical care (specifically, where patients were referred for major cardiovascular surgical procedures and for neurosurgery). Each HRR is composed of approximately 10 HSAs. Our baseline leave-out costs instrument was created using un-covered costs of individuals outside of the cross-border HSA but within the state. We create an alternative leave-out costs measure based on individuals outside of the HRR of the cross-border HSA but still within the state. This restricts us to using variation that is geographically further away of the HSA of interest.

Table J2 reproduces the baseline utilization and spending regressions using this alternative instrument definition. The implied Medigap effects are broadly similar to those estimated using our baseline instrument. For instance, using this alternative instrument, we find that Medigap raises total Medicare payments by $1,822, which is very similar to the $1,396 baseline effect (Table

[4]The sample used for this specification drops individuals in cross-border HSAs that reside more than 30 km from the border based on ZIP code centroid. These specifications still include all individuals who do not reside in cross-border HSAs, as these individuals continue to assist in identifying the coefficients of the control variables.

[5]Within cross-border HSAs, the mean distance from a ZIP code centroid (our most disaggregated measure of location) to the state boundary is 25 km and the median distance is 16 km.

Table J1: Robustness Checks

	Leave-Out Costs (Hundreds)			Mean of	Implied Medigap Effect	
	Est	Std. Err.	P-Value	Dep. Var.	Level	%
Baseline Specification						
Medicare Payments	-67.02	(33.11)	0.043	6,291	1,396.25	22.2%
Alternative Specifications (Dep Var is Medicare Spending)						
Census ZIP Code-Level Controls Included	-59.96	(30.16)	0.047	6,291	1,249.09	19.9%
Region-Year Fixed Effects Included	-55.54	(31.74)	0.085	6,291	1,157.02	18.4%
Unaffected Procedures						
Urgent RVUs (Clemens & Gottlieb Def'n)	5.44E-02	(6.76E-02)	0.421	4.274	-1.13	-26.5%
Urgent Admissions (Card, Dobkin, & Maestas Def'n 1)	-1.31E-03	(1.03E-03)	0.201	0.077	0.03	35.4%
Urgent Admissions (Card, Dobkin, & Maestas Def'n 2)	-6.89E-04	(1.03E-03)	0.505	0.125	0.01	11.5%
Unaffected Individuals						
Non-Elderly Adults in NHIS						
Hospital Days	0.042	(0.048)	0.378	0.342	-0.88	-257.1%
Hospital Stays	0.011	(0.006)	0.060	0.086	-0.23	-266.2%
Physician Office Visits (Indicator for ≥ 2)	0.019	(0.012)	0.106	0.528	-0.39	-74.6%
Self-Reported Health	0.040	(0.042)	0.340	1.999	-0.83	-41.4%
Robustness to Spatial Trends (Dep Var is Medicare Spending)						
Restricted to ZIP Codes Within 30 km of Border	-70.62	(32.25)	0.029	6,291	1471.15	23.4%
Placebo Borders	-14.27	(22.97)	0.535	6,291	297.23	4.7%

Notes: Table displays estimates from regressions of spending and utilization measures on leave-out costs, HSA fixed effects, and controls for age, sex, race, health risk, and GAF/OWI adjustment factors. Each row displays the results from a separate regression. The implied Medigap effect is calculated by dividing the estimate by the coefficient on leave-out costs from the baseline demand specification. This analysis draws on data from the pooled 1999-2005 CMS Beneficiary Summary File, CMS Denominator File, CMS Carrier File ("Urgent RVU" analysis), NHIS ("Unaffected Individuals" analysis), and CMS MedPAR ("Urgent Admissions" analysis). Aside from the NHIS, for each of these datasets we use a sample definition analogous to the baseline sample described in Panel B of Table 2. The "Unaffected Individuals" analysis utilizing the NHIS data focuses on the sample of non-elderly adults, excluding those with Medicare coverage. Standard errors are clustered at HSA level except for the "Placebo Borders" specification in which standard errors are clustered at the HSA-state level (see text for a full description). Dollar values are inflation-adjusted to 2005 using the CPI-U.

7). Note that while the alternative instrument reduces concerns about bias from spatial correlation, dropping these individuals reduces the power of our instrument, inflating the standard errors.

Table J2: Robustness: Alternative Instrument Definition

Dependent Variable	Leave-Out Costs (Hundreds) Est	Std. Err.	P-Value	Mean of Dep. Var.	Implied Medigap Effect Level	%	N
Medicare Payments							
HSA defn	-67.0	(33.1)	0.043	6,291	1,396	22.2%	130,895,953
HRR defn	-87.4	(68.6)	0.203	2,648	1,822	68.8%	130,895,953
Part B Payments							
HSA defn	-21.8	(15.5)	0.159	2,648	454	17.2%	130,895,953
HRR defn	-41.0	(24.8)	0.097	2,648	855	32.3%	130,895,953
Part A Payments							
HSA defn	-47.6	(22.8)	0.037	3,021	992	32.8%	130,895,953
HRR defn	-43.9	(48.4)	0.365	3,021	915	30.3%	130,895,953
Total RVU							
HSA defn	-1.29	(0.50)	0.009	70.77	26.88	38.0%	23,708,295
HRR defn	-1.85	(1.16)	0.112	70.77	38.50	54.4%	23,708,295
Part B Events							
HSA defn	-0.418	(0.181)	0.021	27.59	8.71	31.6%	130,895,953
HRR defn	-0.859	(0.337)	0.011	27.59	17.89	64.8%	130,895,953
Part A Days							
HSA defn	-0.062	(0.019)	0.001	2.21	1.29	58.6%	130,895,953
HRR defn	-0.081	(0.046)	0.079	2.21	1.69	76.6%	130,895,953

Notes: Table displays estimates from regressions of Medicare payments on leave-out costs, HSA fixed effects, and controls for age, sex, race, health risk, and GAF/OWI adjustment factors (see Section II, Equation 4). Each row displays the results from a separate regression. The rows indicate whether the leave-out costs instrument is defined at the HSA or HRR level. The first stage coefficient from a specification with the HRR-level instrument and HSA fixed effects is 0.19 and thus we scale the HRR leave-out costs coefficient estimates by 5.26 to make them comparable to the HSA-level estimates, which have first-stage of 0.94 to 1.1 across specifications. The implied Medigap effect is calculated by dividing the estimate by the coefficient on leave-out costs from the baseline demand specification. This analysis draws on data from the pooled 1999-2005 CMS Beneficiary Summary File and CMS Denominator File. This analysis uses the baseline sample described in Panel B of Table 2 (N=130,895,953). All dependent variables are top-coded at $64,000. Standard errors are clustered at the HSA level. Dollar values are inflation-adjusted to 2005 using the CPI-U.

K Robustness of Policy Counterfactuals

Table K1 examines the sensitivity of our estimates of the budgetary effect of a tax on Medigap premiums. To examine robustness to heterogeneity in the price-elasticity of demand, rows of Table K1 re-calculate the effect of a 15% tax using the different demand estimates from Table 5. We find that across these different estimates, the total budgetary savings to Medicare range from 3.9% to 4.8%. We also show standard errors for each specification. For our baseline estimate of 4.3% total savings, the standard error is 1.7 percentage points.

Table K1: Tax Counterfactuals: Robustness to Alternative Demand Estimates

Tax	Demand Parameter Used	Medigap Market Share	Tax Revenue (per Beneficiary)		Medicare Savings (per Beneficiary)		Total Budgetary Impact (per Beneficiary)			
			Estimate	SE	Estimate	SE	Estimate ($)	Estimate (%)	SE ($)	SE (%)
15%	Supp Cov, Combined Hsa (baseline)	35%	$94	$16	$179	$88	$273	4.3%	$105	1.7%
15%	Supp Cov, MCBS Hsa	30%	$81	$27	$179	$88	$260	4.1%	$115	1.8%
15%	Supp Cov, NHIS Hsa	40%	$106	$19	$179	$88	$285	4.5%	$107	1.7%
15%	Medigap, MCBS Hsa	26%	$69	$43	$179	$88	$248	3.9%	$131	2.1%
15%	Supp Cov, Combined Hrr	37%	$100	$11	$179	$88	$279	4.4%	$100	1.6%
15%	Supp Cov, MCBS Hrr	29%	$77	$20	$179	$88	$256	4.1%	$108	1.7%
15%	Supp Cov, NHIS Hrr	45%	$121	$11	$179	$88	$300	4.8%	$100	1.6%
15%	Medigap, MCBS Hrr	24%	$64	$35	$179	$88	$243	3.9%	$123	2.0%

Notes: The first column lists the tax as a percentage of the $1,779 average Medigap premium. The second column describes the demand estimate from Table 5 that is used in the calculation, assuming full pass-through of the tax. The linear demand curve used in these calculations has a slope equal to $\partial q_{ijk}/\partial$Leave-Out costs$_{jk}$ (as the coefficient on leave-out costs in the premium regressions is approximately one) and an intercept pinned down by the equilibrium average price and quantity (p=1,779 and q=0.48). The remaining columns list the tax revenue, cost savings from Medigap dis-enrollment, and total budgetary impact, respectively. These results are based on the estimated $1,396 Medigap externality. To calculate the standard error on the total budgetary savings, we first separately calculate the standard error on the tax revenue (from the corresponding demand estimate) and the standard error from the Medicare cost savings from Medigap dis-enrollment (from the reduced form cost estimates). We then obtain the standard error on the total savings using the Delta Method assuming no covariance between the demand and cost estimates. Dollar values are inflation-adjusted to 2005 using the CPI-U.

References

Card, David, Carlos Dobkin, and Nicole Maestas. 2009. "Does Medicare Save Lives?" *Quarterly Journal of Economics*, 124(2): 597–636.

Clemens, Jeffrey, and Joshua Gottlieb. 2014. "Do Physicians' Financial Incentives Affect Treatment Patterns and Patient Health?" *American Economic Review*, 104(4): 1320–1349.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2016. "Sources of Geographic Variation in Health Care: Evidence from Patient Migration." *Qquarterly Journal of Economics*, 131(4): 1681–1726.

GAO. 2001. "Retiree Health Insurance: Gaps in Coverage and Availability." Statement of William J. Scanlon Director, Health Care Issues: Testimony Before the Subcommittee on Employer-Employee Relations, Committee on Education and the Workforce, House of Representatives. Government Accountability Office.

Wennberg, John E., Elliott S. Fisher, and Jonathan S. Skinner. 2002. "Geography and the Debate over Medicare Reform." *Health Affairs*, 21(2).