The Baby Boomer Generation is Reaching Retirement Age, How Has the Growing Elderly Population Effected the Costs of Household Services?

By Ian Rolls

The U.S. Census Bureau predicts the number of people ages 65+ to increase from 35 million people in 2000 to 85 million in 2050. Given that these individuals are more likely to acquire disabilities as they age, they will need people to provide household services that they can no longer preform themselves. This study aims to find relationship between the percentage of people in a country over age 65+ with the cost of household services. Although the baby boom is unique to the United States and Canada, this study examines seven countries worldwide in a longitudinal panel analysis. This study finds that there is no statistically significant effect of aging on the price of household services. (JEL I10 G50 J14)

In mid-twentieth century United States, there was a large baby boom, which is a period marking a significant increase in the birth rate. The average marriage age decreased from 22 to 20 and, between 1955 and 1964 alone, around 42 million babies were born.[1] The generation born between the years 1946 and 1964 are now labeled together as a generation, as the "baby boomers." Because these birth rates dropped after 1964, the population increase did not continue and now baby boomers make up 21.19% of the current U.S. population, only .8 percentage points behind the second largest, the millennials. This is unusually high given the world's trend towards continually higher birth rates. As these baby boomers grow older, they will require household services to fill in gaps caused by naturally occurring aging disabilities. However, if the demand for these services increases from the large number of baby boomers, will the cost rise as well? Regardless of how it effects the cost through the demand side, there is also a supply side effect of increased aging. Because the household service is a low skill occupation, it is a low paying job with a median wage of only $25,280 per year. [2] For the labor supply to increase, the median wage would have to rise. This study seeks to understand the relationship between the cost of household services and the amount of people above age 65. This will provide insight to the rising magnitude of costs faced by aging baby boomers and could point towards some policy measures to alleviate financial pressure on the elderly. I opt to use both an ordinary least squares (OLS) model and a longitudinal panel model to examine a correlation between these variables.

I. Literature Review

Although there is a robust literature base describing the relationships between aging and healthcare spending, there is not much research on the correlation of aging and household service costs. However, there is overlap between because household services are a subset of healthcare spending. Uwe E. Reinhardt explores this first relationship between aging and the demand for healthcare in his study, using individual level panel data from the Medical Expenditure Panel Survey (MEPS). With the MEPS data, he concludes that the effect of aging on health care demand is negligible, that the increasing demand for healthcare is due to other factors. Reinhardt

[1] B. Robey, "A Guide to the Baby Boom," *The American College of Cardiology Extended Learning Supplement* (September 1982) 16–21.
[2] Bureau of Labor Statistics, "Home Health Aides and Personal Care Aides: Occupational Outlook Handbook," https://www.bls.gov/ooh/healthcare/home-health-aides-and-personal-care-aides.htm.

argues age factors alone would make average annual per capita health spending grow at a compound growth "rate of only 0.4 percent" per year.[3] He cites other studies that predict healthcare spending has a compound growth rate over the period [1990-2030] of 8.3 percent.[4] He argues a large majority of healthcare demand growth is due to other factors. In reality, the increased longevity of lifespan due to medical technologies will also increase the time that people need to spend in more intensive home care facilities, and therefore it is important to examine a specific aspect of healthcare. Furthermore, Reinhardt only examines the effects that the increasing population has on healthcare demand but doesn't consider the change in supply from the decreasing population proportion of people age 15-64. He argues that the increased supply side costs will be counteracted by new technology that can streamline household services.

Additionally, we should examine the trends in the household services industry. Johnson, Toohey and Wiener study how we should meet rising aging and the need for long term care. Within this study, they examine the trends towards assistance, claiming that between 2000 and 2040 amount of adults ages 65+ with disabilities will more than double, increasing from 10 million to 21 million. Furthermore, the number of frail older adults relative to the number of adults ages 25-64 will increase from 6.6 to 10.6 percent.[5] Redfoot Et al explains how the boomer population is expected to increase by 79% at 2030 while the population of potential caregivers aged 25-64 is expected to only increase by 1%.[6] This will put a financial constraint on the younger generation that could only be ameliorated with higher wages or technological improvements. Besides having less people to care for the elderly, the amount of educational gains among adult women have increased, meaning there is a higher opportunity cost for providing care by boosting their potential earnings.[7] Families can circumvent the high costs of formal care if a family member assumes the role of caregiver, however this would not be as likely given the increased educational attainment for women, meaning more elderly might need to pay for formal caregiving.

One crucial aspect determining the future cost of health care is the disability rate, which has been trending downwards although experts are unsure of how it will continue.[8] Cutler and Sheiner examine disability trends dropping ~1% per year, they argue that falling disability rates will offset the cost burden brought by increasing aged population.[9] However, Johnson Et al calculated in their best-case scenario where disability continued to decrease even farther than the projected 1% per year, there will still be large burdens on families and institutions. In their intermediate scenario where the disability rate remained at its moderate yearly decrease, they predict the number of total paid hours of care to triple, meaning financial planning becomes necessary to prepare for their long-term needs.

---

[3] Reinhardt, 33.
[4] Ibid, 29.
[5]Richard W Johnson, Desmond Toohey, and Joshua M Wiener, "Meeting the Long-Term Care Needs of the Baby Boomers," *The Retirement Project*, 13.
[6] Donald Redfoot, Lynn Feinberg, and Ari Houser, "AARP Public Policy Institute," 1990, 5.
[7] Ibid, 18.
[8] Ibid, 8.
[9]David M. Cutler, and Louise Sheiner, "Demographics and Medical Care Spending: Standard and Non-Standard Effects," *Finance and Economics Discussion Series*, Board of Governors of the Federal Reserve System (U.S.), 1999, https://ideas.repec.org/p/fip/fedgfe/1999-20.html, 43.

Houser Et al examine trends in assistive technology use insides homes with older people with disabilities. They found that around 60% of people with only 1-2 difficulties with activities of daily living (ADL) and 22% of those with 3-4 ADLs only rely on assistive equipment to help perform these activities. Both Reinhardt and Johnson Et al point towards technology development as a solvency method to increasing household service costs and decreasing labor supply. This could explain why we aren't currently seeing large price increases as some of the earliest boomers are reaching retirement. If more technology is produced to alleviate difficulties with disabilities, this could decrease the demand for labor and keep costs steady.

II. Data and Descriptive Statistics

As this study uses a longitudinal panel model, I use years as time, I have 21 years for each state, spanning from 1998 to 2019, although some of my control variables only were available from 2000-2018. This causes those extra three years of observations to be removed in certain models. This study examines the effects of an increased elderly population on the price of household services in multiple countries, therefore I chose 7 countries to obtain data from: Australia, Canada, Denmark, France, Switzerland, The United States, and The United Kingdom. Because I use a longitudinal panel model, I collected data for each year within each state, a total of 143 observations (Denmark didn't have data for all years). The model's dependent variable is the consumer price index (CPI) of household services for each country respectively divided by the total CPI of the country. CPI is especially helpful for this data because it is a measure of prices not related to currency. Even though each country uses different currencies, CPI is a universal measure of prices. However, CPI alone is not a good measure for prices because CPI does not control for inflation. To create a real measure of prices I divided the household service specific CPI by the holistic CPI to isolate the changes in price not attributed to inflation. All the CPI measures have a base year, the base year is set to 100.0, and fluctuations from that number measure the change in prices relative to the base year. Although each country used a different base year for their CPI, I standardized all the CPI's to have a base year of 2015, meaning the ratio of prices for industry to country is 1.0 for every country in 2015.

My independent variable is the percentage of a population above age 65, I use World Bank data in my dataset. I specifically use the percentage of population above age 65 instead of the number of people above 65 because the percentage allows me to include the supply side effect of the population increase. As the percentage of people over age 65 increases, the percentage of people under age 65 decreases, capturing both effects. Reinhardt's dataset uses population growth instead of change in the percentage of population ages 65+. This measure cannot account for a drop in the amount of potential caregivers ages 25-64, and therefore doesn't examine the supply side effect of household service prices. He admits that his study cannot capture the supply side effect, but he argues that technological advances will fill in for workers to stabilize the cost.

I also control for some factors that might explain the variation in my dependent variable. I use data from the Organization for Economic Co-operation and Development (OECD) for these controls because they have a wide variety of useful health data for many countries. I control for healthcare spending per capita, healthcare spending as a percentage of GDP, and the number of

long-term care recipients. Lastly, I create a binary outcome labeled highcost which is 1 if greater than 1.0 and 0 if less than 1.0, because at value 1.0 the ratio of household services CPI to country CPI is equal, meaning the increase in price matches inflation.

III. Econometric Model and Results

Here is the basic econometric model, an OLS model with response variable CPI ratio, which is the ratio of the household services CPI to the country CPI, and the explanatory variable being the percentage of population above age 65

$$CPIratio = \beta_0 + \beta_1(Percent\ of\ Population\ above\ age\ 65) + \beta_2(Healthcare\ spending\ as\ a\ percentage\ of\ GDP) + \beta_3(Healthcare\ Spending\ per\ capita)$$

Next is the longitudinal panel model with the same explanatory and response variables as with the OLS model, however these variables are separated into time and state specific responses, with t being the time periods (1998-2019) and t being the states (Australia, Canada, Denmark, France, Switzerland, UK, U.S.). Each state and time permutation are contained.

$$CPIratio = \beta_0 + \beta_1(Percent\ of\ Population\ above\ age\ 65)_{it} + \beta_2(Healthcare\ spending\ as\ a\ percentage\ of\ GDP)_{it} + \beta_3(Healthcare\ Spending\ per\ capita)_{it} + \beta_4(Healthcare\ spending\ as\ a\ percentage\ of\ GDP \times\ Percent\ of\ Population\ above\ age\ 65)_{it}$$

I chose to look at multiple countries even though the baby boom is specific to the United States. I wanted to use a longitudinal panel model because I could ignore year fixed effects to get a more consistent estimate of my β_1 value. Furthermore, examining different countries with different healthcare systems would allow me to get a more universal estimate for the effect of aging on cost.Before analyzing any of the data, I predicted the marginal effect of my explanatory variable to be positive, because I assumed as more people enter a market with a shortage of labor, the prices would have to increase to become competitive enough for more firms and laborers to enter the market. Furthermore, I assumed the large increase in the population ages 65+ to offset any disability decrease because people are living longer and might require more care.

However, my results show that the effect of aging on the costs of household services is negative, with a $\hat{\beta}_1$ value of -.123, (See Column 3) meaning, for every 1 percentage point increase in the percent of people over age 65, the ratio of the industry price to the country price decreases by -.123. In other words, prices of household services are predicted to decrease by 12.3 percent. This is of a significant magnitude, one that I never expected because my research deduced there would either be a significant increase in costs or no change, although multiple studies predicted technological advances would decrease prices, which might account for the high negative number. I jointly tested my coefficients in each model, and in the panel regression with controls they are jointly significant at the 10% level. However, the panel model without controls was statistically insignificant. One endogeneity concern that might have caused this

result was omitted variable bias. One datapoint that is highly correlated with both my explanatory and response variable is the disability rate. However, I would have to go to each individual country's national databank to find that data and I most likely would only find it for 4 of the 7 countries, so I wasn't able to include it in this dataset. This error term is likely endogenous in my explanatory variable and skewing my results. My $\hat{\beta}_1$ value is most likely an overestimate (too negative) because if $\hat{\beta}_1 = \beta_1 + \hat{\alpha}_2\gamma_{12}$ and in this scenario γ_{12} is the correlation between the disability rate and the percentage of population over 65 years, $\hat{\beta}_1$is my predicted marginal effect, and $\beta_1$is the true marginal effect. Because the correlation between disability rates and the percentage of people positive over age 65 is negative (currently as the boomer population is increasing disabilities are decreasing), and the marginal effect of disability rates is positive (the more disabilities the higher the cost), the term right of the plus sign would be negative, meaning my true marginal effect is smaller than the predicted marginal effect. Because the decreasing disability rate has been decreasing costs, we cannot isolate the true effect of increasing population portion on the costs of household services.

Table 1: Effects on the CPI ratio (household service industry CPI/Country CPI)

VARIABLES	(1) Pooled OLS	(2) Panel without controls	(3) Panel with controls	(4) Logit Regression
Percent Pop Above 65	-0.0103***	-0.00863	-0.123**	-0.799***
	(0.00392)	(0.0157)	(0.0468)	(0.0903)
Heathcare Spending (% of GDP)			-0.118*	-0.948***
			(0.0548)	(0.172)
(Healthcare Spending)^2			0.000706	-0.00346
			(0.00156)	(0.00366)
Healthcare Spending (per capita)			-2.35e-05**	4.43e-05**
			(6.79e-06)	(2.02e-05)
Interaction Term			0.00989**	0.0666***
			(0.00398)	(0.00885)
Constant	1.170***	1.145***	2.583***	
	(0.0663)	(0.244)	(0.635)	
Observations	143	143	123	132
R-squared	0.068	0.035	0.565	
Country Fixed Effects	NO	NO	YES	YES
F-stat	6.838	0.304	3.856	
Adjusted R-Squared	0.0616	0.0285	0.546	
Number of Countries		7	7	
Hausman p-value			0.0153	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Pooled OLS model in column (1) while the panel data is in columns (2), (3) and the logit regression is in (4). The interaction term is (Healthcare Spending % of GDP * Percent Pop above 65). The response variable for the logit regression is "highprice" which is a measure of the CPI ratio that equals 1 if the CPI ratio > 1.0 and equals 0 if the CPI ratio is ≤ 1.0. There are only 123 observations in the panel model with controls because the data for the controls started at year 2000 instead of year 1998. Although Healthcare Spending^2 is not statistically significant, it is jointly

significant with Healthcare Spending which is why I included it in the table. The Hausman test points towards a fixed effects model to be a more consistent estimator.

Another endogeneity problem my regression could suffer from is sample selection bias. Because I collected the price data from each individual country's national data bank, some countries may have different expenditures that might fall under household services. None of the national data banks included a detailed summary of each category, only an example, which made data collection confusing. For example, the United States BLS labeled household services as "personal care," while France and Switzerland put haircuts and beauty expenditure under "personal care," there could be expenditures that fit outside the realm of household services in the dataset that I can't account for.

One way I could solve my omitted variable bias endogeneity problem would be to create instrumental variables, these variables would need to be highly correlated with my explanatory variable but uncorrelated with my response variable, this way I could capture the causality in the explanatory variable without having endogenous error. One instrumental variable I thought of could be the number of hearing aids sold in a year, because this would be highly correlated with the percentage of people over age 65, but I don't think it would be correlated with the cost of household services because those two markets are mostly separate.

I also created an interaction term between the percentage of the population above age 65 and healthcare spending as a percent of GDP to examine the effect of healthcare spending on the price of housing services given different percentages of the population over age 65. The coefficient for the interaction term is 0.00989, meaning the marginal effect of healthcare spending as a percentage of GDP is (0.00989*Percent Pop above 65). The marginal effect changes based on how large this percentage is.

Table 2

VARIABLES	(1) Logit Regression	(2) Binary OLS
Percent Pop Above 65	-0.799***	-0.782***
	(0.0903)	(0.0794)
Healthcare Spending (% GDP)	-0.948***	-0.839***
	(0.172)	(0.163)
(Healthcare Spending)^2	-0.00346	-0.00706*
	(0.00366)	(0.00395)
Healthcare Spending (per cap)	4.43e-05**	7.38e-05***
	(2.02e-05)	(2.78e-05)
Interaction Term	0.0666***	0.0642***
	(0.00885)	(0.00804)
Constant		11.47***
		(1.310)
Observations	132	132
R-squared		0.423
Country Fixed Effects	YES	YES
F-stat		45.26
Adjusted R-Squared		0.400

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Logit regression in (1) is the same as the previous logit regression in (4) of table, it is repeated to put it in side-by-side comparison with the binary OLS model. Both models use the response variable "highprice."

Lastly the logit and binary OLS model show a similar result to the panel regression, but instead they are measured with the highprice variable. Both outcomes are statistically significant with a high magnitude, with a marginal effect of -0.799 for the logit regression and -0.782 for the linear probability model (Binary OLS). For the logit regression, this means, for a one percentage point increase in the percent population above age 65, the probability of a price increase decreases by 79.9%. Likewise, with the linear probability model, the probability of a price increase decreases by 78.2%.

I focused on a different aspect of the topic that my sources did, as I could not find much economic literature that focused on costs. Although Reinhardt is an economics professor at Princeton, he opted to only use graphs instead of tables, which are still helpful. The first table

Table A5. Determinants of Paid Long-Term Care Services, for Frail Older Adults with Surviving Children

	Any Paid Home Care		Any Nursing Home Care	
	Odds Ratio	S.E.	Odds Ratio	S.E
Hourly Price of Children's Time	1.07***	0.02	1.08***	0.02
Number of Children				
Daughters up to 3	1.10	0.08	0.97	0.10
Sons up to 3	1.03	0.07	1.13	0.11
Indicator for 4 or more daughters	0.99	0.28	2.06*	0.78
Indicator for 4 or more sons	0.85	0.21	0.46**	0.18
Age	1.08***	0.01	1.05***	0.01
Severely Disabled	12.20***	1.59	10.16***	1.80
Male	0.81	0.12	1.03	0.21
Race and Ethnicity				
African American	1.02	0.22	0.65	0.19
Hispanic	1.11	0.30	0.61	0.28
[Ref: White or other]	…	…	…	…
Marital Status				
[Ref: Widowed]	…	…	…	…
Divorced or separated	1.13	0.28	0.99	0.30
Never married	2.26	1.47	…	…
Currently married	0.35***	0.06	0.27***	0.07
Married to Spouse with Disabilities	1.82***	0.40	1.34	0.43
Education				
Not high school graduate	1.04	0.16	1.10	0.21
[Ref: High school grad]	…	…	…	…
Four or more years of college	1.11	0.24	0.78	0.22
Ratio of Income to Poverty Level				
No more than 1	1.70*	0.48	2.05*	0.76
1.01 to 2	1.41	0.34	1.69	0.55
2.01 to 3	1.21	0.31	1.35	0.46
3.01 to 4	…	…	…	…
Greater than 4	1.65*	0.44	1.86*	0.64
Wald chi-square statistic	519.70		332.31	

Source: Authors' computations from the 2002 Health and Retirement Study (HRS).
Notes: Estimates were from logit models of any unpaid help from adult biological children, any unpaid help from other sources, any paid home care, and any nursing home. Help received at home was measured during the month preceding the survey interview, and nursing home care was measured at the time of the interview. S.E. denotes the standard error of the odds ratio. The sample was restricted to 2,713 adults ages 65 and older with at least one limitation with an activity of daily living (ADL) or instrumental activity of daily living (IADL) and with surviving children. The price of children's time was set equal to the imputed hourly wage of the child with the lowest wage.

* $.05 \leq p < .1$ ** $.01 \leq p < .05$ *** $p < .01$