Co-Residence Among Young Adults and Their Parents:

Insight from the PSID

Rumian R. Reza

Econometric Analysis of Panel Data 5655

Dr. Myungsup Kim

Fall 2021

University of North Texas

December 4, 2021

1. Introduction

I define *co-residence* as the sharing of households between the generation of young adults (labeled as G1) and their parents, members of the older generation (labeled as G0)[1]. The familiar Overlapping Generations Model (OLG) designated the interaction between G0 and G1 as a macroeconomic question about intergenerational transfers (Diamond, 1965). Here, however, I use microeconomic panel data during a significant economic shock (the 2008 financial crisis and its precursors) to identify correlates of co-residence[2]. I focus on variables within the nuclear family and employ linear and nonlinear binary response models. Using the Panel Study of Income Dynamics (PSID), I find that Medicare use by G0 is positively correlated with co-residence probability. This finding confirms implications that follow from Rust & Phelan (1997). Furthermore, co-residence probability is significantly higher if the G1 individual's mother is unmarried, confirming the study by Pilkauskas & Cross (2018). The impact of the Great Recession is undetermined; a relevant, well-defined measure is needed for further insight. Below, Section 2 discusses the background of my empirical expectations, Section 3 describes the data and outcome variable, Section 4 is about the estimation strategies, Section 5 briefly discusses the results, and Section 6 mentions pending analysis.

2. Literature Review

Demographic literature documents increasing co-residence over recent decades. One claim is that G1 is growing dependent on G0: housing wealth is concentrated among older individuals (Lusardi & Mitchell, 2007; Zhao & Burge, 2017). Another claim is that longer life

---

[1] In other words, co-residence is the state of G0 and G1 living together (with no claims regarding dependency).
[2] That is, I take G1's perspective in this analysis to offer insight on motivations and deterrents of co-residence and generational living arrangements.

expectancy has G0 growing dependent on G1, particularly in terms of long term care and retirement insecurity (Brown & Finkelstein, 2011; Poterba, 2014; Rust & Phelan, 1997). Trends in multigenerational household living arrangements in the United States have also been documented in the sociological literature[3] and, to a lesser extent, in the macroeconomic literature[4]. In general, co-residence rates among young adults and their older family members have been increasing in recent decades (Pilkauskas et al., 2020). But empirical evidence is desired regarding motivations and deterrents of co-residence. There is a question of the generational perspective from which one should analyze the decision (G0 or G1). Bartczak et al. (2021) emphasize the generational perspective of young adults (G1) as they are the "squeezed middle," with downward economic pressure from G0 and upward pressure from G2 (G1's children, also known as G0's grandchildren).

Evidence suggests that generationally independent[5] living arrangements are a normal good (Engelhardt et al., 2005; Pilkauskas & Michelmore, 2019). In this sense, an increase in generation-specific household resources such as Medicare should be associated with a lower chance of co-residence, ceteris paribus; Pilkauskas & Cross (2018) pose this as a possibility. On the other hand, Rust & Phelan (1997) find that prospective retirees prolong retirement[6] until they are Medicare-eligible (to avoid gaps in healthcare coverage). Upon retiring, G0 may further

[3] For instance, Wiemers et al. (2017) find that factors related to the mothers of adult children are the main determinants of co-residence. Co-residence trends have also been shown by sociological literature to extend beyond the nuclear family (Bengtson, 2001; Harvey et al., 2021). My study borrows much insight from Pilkauskas & Cross (2018) and Pilkauskas et al. (2020).

[4] See Altonji et al., (1989), Diamond (1965), Hayashi et al. (1996) and Kotlikoff & Morris (2008).

[5] "Generationally independent" refers to G1 and G0 living independently; G2, if existent, resides separately from G0 (though *with* G1).

[6] Note that Engelhardt et al., (2020) find that early SS-claiming behavior is associated with a higher poverty rate among the elderly, suggesting a higher likelihood of co-residence for those who withdraw SS funds before the age of 65.

reduce risk by pooling assets (such as housing) with their adult children. The impact of G0 Medicare coverage can therefore be argued as theoretically ambiguous.

Another factor identified by Pilkauskas & Cross (2018) is the relationship status of a young adult's mother: co-residence is less likely if the G0 mother is married[7]. Also, children who have an immigrant parent[8] might be likelier to share the household with their parent(s) as adults (Gubernskaya & Tang, 2017). Furthermore, a G1 individual having a child (alternatively, whether G0 has grandparental status) may increase G0's tendency to transfer childcare time to G1. This, in turn, could motivate co-residence among a G0 mother and, for instance, an unmarried G1 parent, though reasons could include unobserved preferences (Rupert & Zanella, 2018). Other factors of G1's co-residence status are G1's age and G0's housing status (or housing wealth): economic independence among young adults may increase when entering prime-working years (Kahn et al., 2013), whereas G0 housing wealth (or homeownership status) could lengthen co-residence duration (Engelhardt, 2008; Lusardi & Mitchell, 2007; Zhao & Burge, 2017). Lastly, both state and national economic trends (such as labor or housing market conditions) may affect co-residence rates[9].

A variable of interest is the level of exposure to the Great Recession. The assessment of its impact on co-residence is mixed. Gustman et al. (2010) claim that economic independence among those near retirement age declined due to the possession of retirement assets tied to stock market investments. De Bresser et al. (2021) extend its deleterious effects to pension wealth among older, higher-income individuals, which could imply a higher co-residence probability.

---

[7] The number of siblings belonging to a young adult is also a potential co-residence determinant (Pilkauskas et al., 2020). However, A time-varying measure of the number of siblings belonging to an individual is not well defined within the PSID, though such a measure can be formed with some effort.
[8] "Immigrant status" is not directly measured in the PSID, but birthplace is (discontinuously).
[9] State-level unemployment rates and regional cost-of-living indices could capture these market dynamics.

Meanwhile, Kahn et al. (2013) find a near-50-percent increase in co-residence rates among G1 adults and their G0 parents over 2000-2010, positing that tumultuous post-Great Recession labor and housing markets were the culprits of this trend. On the contrary, Pilkauskas & Cross (2018) find that this increase in co-residence rates started before the Great Recession and preceded the housing market collapse as well. The conflicting interpretations of co-residence rates over this time period motivates a closer, microeconomic look.

3. Data Source

3.1. Panel Study of Income Dynamics (PSID)

I use the biennial waves of the PSID from 2003 to 2011, inclusive[10]. The PSID follows individuals and family units[11] for an indefinite duration. New sample members enter the PSID through connections with existing sample members (via marriage or birth, for instance)[12]. Data is available at the individual level and the family level, with family Heads[13] and spouses of family Heads intensively measured at the family level[14]. I rely on the PSID's methodology for this study: children in the PSID may age and "split off" from the original nuclear family as young adults, effectively becoming autonomous economic entities (forming family units of their own). These young adults are then distinctly surveyed as new family units (as in, the young adults are now independent family Heads themselves). Splitting off from a nuclear family unit is one

[10] Link to data (valid as of December 2021): https://simba.isr.umich.edu/data/data.aspx

[11] A family unit is a unit of analysis within the PSID and consists of a "family Head" (presumably the primary household resource allocator), the family Head's spouse, and dependents, if any. Family units are distinctly defined from households in that one household can contain more than one family unit; this methodology forms the basis of my analysis.

[12] Other ways to enter the PSID include via supplements or immigrant refresher survey waves.

[13] A family Head (also called a Reference Person in recent PSID waves) is the main (representative) respondent to family-level survey questions. Supposedly, a family Head could be designated as the primary resource allocator in the family unit. PSID documentation offers rigorous definitions.

[14] If the sample were restricted to family Heads and spouses of Heads only, then some family-level measurements can have individual-level interpretations.

mechanism for entering the PSID as an informative sample member[15]. In this regard, the PSID's methodology is convenient in that young adults who split off from the nuclear family but move back in with the nuclear family are still recognized as belonging to separate family units *despite* sharing the household. The fact that the PSID is a survey of family units rather than households is crucial for the nature of my study.

I link generations together by using the PSID's Family Identification Mapping System (FIMS). FIMS is available as a retrospective (or prospective) intergenerational map, and it links individuals with their parents (or children) in preceding (or subsequent) generations. Both maps can account for different generational depths[16]. In my case, the retrospective FIMS allows me to observe individuals and their parents[17] simultaneously as one observation. I then merge data pertaining to individuals and their parents for the same set of variables to form a cross-year individual-level data set with the G1 young adults as the cross-sectional units of analysis. However, when an individual is *not* a family Head or spouse of a family Head, responses to family-level survey questions require careful interpretation.

Other sources of data include the U.S. Bureau of Labor Statistics' (BLS) state-level seasonally adjusted unemployment rates, (aggregated by FRED[18]), the BLS' seasonally adjusted consumer price index[19], and the state-level Freddie Mac Housing Price Index[20] (FMHPI[21]).

---

[15] An informative sample member is one that is the main respondent to family-level survey questions. Another way to be recognized as a new informative sample member is by divorcing from an existing Head, then living independently.
[16] For example, I can link young adults, their parents, and their parents' parents. I would therefore observe three generations as one observation (a depth of three generations).
[17] This includes both biological and adoptive parents of a G0 individual, though I analyze individuals who do not have adoptive parents in the PSID (adopted individuals are dropped from the sample).
[18] Link to data (valid as of December 2021): https://fred.stlouisfed.org/release?rid=112
[19] Link to data (valid as of December 2021): https://fred.stlouisfed.org/series/CPIAUCSL
[20] Link to data (valid as of December 2021): http://www.freddiemac.com/research/indices/house-price-index.page
[21] During estimation, I found FMHPI to be insignificant across all models. Thus, the housing price index does not appear in the discussion below.

3.2. Variables

Although Pilkauskas & Cross (2018) make a case for including SS reception as a variable, SS reception is a function of G0's age; the same is true for Medicare reception. I instead use G0's Medicare reception as a variable and consider the total non-SS transfer income received by G1's parents (head and spouse combined). Like SS, non-SS transfer income may demonstrate a fixed income stream. This could be positively correlated with the probability of co-residence without (or at least not as heavily) relying on G0's age for transfer eligibility. Both variables (Medicare and non-SS transfer income) are therefore included in the models.

The outcome variable is *co-reside*, an indicator equal to one if a given G1 individual resides with either G0 parent at the time of the interview[22]. An alternative outcome variable of interest is the "Split-off Indicator" which equals one when an individual's family unit is *new* due to splitting off from the original family unit, thereby becoming an economically independent vehicle[23]. The Split-off Indicator appears to be a more promising outcome variable than a contemporaneous measurement of co-residence. Due to survey intensity (biennial), much is unobserved, so a contemporaneous measure of co-residence fails to capture transitory co-residence that might take place in between survey waves. The Split-off Indicator, however, does capture what occurs *in between* the survey waves[24]. Unfortunately, there is insufficient variation within the Splitoff-Indicator variable in the analysis sample. I therefore resort to the contemporaneous measurement of co-residence as my outcome variable. An inspection of the

[22] Specifically, I define *co-reside* as an indicator equal to one if the G1 young adult has the same household ID as their mother, father, or both.
[23] For instance, a young adult who is moving out of the nuclear family within the PSID can be assigned a Split-off Indicator equal to one.
[24] Despite the advantage, this variable may not necessarily capture G1's economic independence from the nuclear family, especially beyond a certain age threshold; other life circumstances (such as divorce) are likelier to warrant splitting off from one's current family unit, and, consequently, the Split-Off Indicator would not be as interpretable as a measure of economic independence *from G0 in particular*.

sample's transition probabilities for *co-reside* reveals that this variable captures G1's initial move-out (a de facto split-off) nonetheless[25].

3.3. Analysis Sample and Summary Statistics

Reported below in Tables 1A and 1B[26] are the (unweighted) summary statistics of a subset of variables. Transition probabilities for a subset of cross-wave time-varying indicators are reported in Table 1C. The cross-sectional unit of analysis is a young adult (aged 18 years or older[27]) who does not have adoptive parents, has at least a mother observed in the PSID[28], is present in the family unit and observed in all five PSID waves over 2003-2011[29], and has no missing values for the variables of interest. Very large values for G0's non-SS transfer income[30] are dropped from the sample as well as negative values[31] for G1's childcare expenditures. These restrictions result in a balanced panel of 933 individuals.

Table 1A shows that G1 individuals have fathers who attain a slightly higher education level on average than mothers, though the standard deviations are quite large. Although the indicators for high school graduation are virtually equal for G0 mothers and fathers, the average

---

[25] The transition probabilities for *co-reside* are reported in Table 1C below.

[26] Table 1B features the summary statistics of the time-varying variables. The averages of time-varying variables can be interpreted as the average individual bearing that value for the variable in any given year. Summary statistics are reportable after the PSID main data is merged with FIMS. The preliminary merge enables the creation of variables that are dependent on intergenerational relationships.

[27] I exclude PSID teens that move out of the household at age 17 (Bastian & Michelmore, 2018). However, the PSID variables that count the number of children within a family unit include anybody within the family unit under the age of 18. Despite the risk of some unobserved split-offs from the nuclear family unit, I impose the over-17 age restriction for the latter reason. Furthermore, co-residing beyond the age of 22 may be considered to be more economically significant than co-residing during ages 18-22. For instance, one could be a college student commuting to school.

[28] This is to capture single-parent households, of which mothers are overrepresented as family heads (Pilkauskas & Michelmore, 2019).

[29] A balanced panel may eliminate attrition bias: G1 members that split off from the nuclear family (headed by G0) may elect to stop appearing in the PSID once they become economically independent.

[30] The PSID warns about very large values for this variable due to nuanced measurement. I drop observations for which non-SS transfer income for either G0 parent exceeds five standard deviations above its mean.

[31] PSID explains that negative values arise for this variable due to the imputation procedure (linear interpolation).

G1 individual has a mother that attains less than a high school diploma (with a large standard deviation noted). Less than half of the units of analysis are female, and a very small proportion are a race other than black or white. Note that not all G0 individuals are unique; there could be G1 siblings in this sample who share parents, in which case G0 individuals are being overcounted. Therefore, these statistics should be interpreted with caution[32].

Table 1B shows that a large proportion of the cross-sectional units of analysis are parents (equivalently, a large proportion of G0 parents are grandparents). Despite this, only 15-percent of G1 individuals in the sample have some positive number of childcare expenditures reported in a given year. This could reflect the fact that childcare expenditures are a family-level measurement, so G1 individuals that have yet to move out of the nuclear family could have this variable measuring their aging parents' childcare expenditures (which are likely zero). Also, note that the average unit of analysis is about 36 years-old, indicating that the average young adult is beyond the typical age of debut economic autonomy (as it is recognized in the United States).

Table 1C pairs well with Table 1B. For example, one can interpret the average of whether G1's mother is married as the proportion of G1 individuals who have married mothers in a given year (between 2003-2011). However, the transition probabilities for this variable reveal cross-wave dynamics. In this case, Table 1C reveals that G1 young adults tend to have married mothers who stay married and unmarried mothers who stay unmarried. Further, note that the 1-to-0 transition probability for whether G1 has one child in the family unit is nearly 30 percent. This can be interpreted as either the G1 individual having a second child or the G1 individual moving out of the nuclear family unit where a younger sibling, for instance, may also reside.

[32] That is, G0 variables, though referring to the older generation, are direct characteristics of G1 rather than G0. Clearly, there is some cross-sectional correlation arising from shared parents.

Regarding the outcome variable, transitioning from co-residing to not co-residing may signal the G1 individual splitting off from the nuclear family unit for the first time or ending transitory co-residence (having already split off earlier in life). A disadvantage of this analysis sample is that a very small proportion of transitions are from the state of not co-residing (living independently, for instance) to co-residing (with the G0 parents). This implies that a majority of those that *do* co-reside at least once during the observed time period are moving *out* rather than moving back in. This could have implications on the signs and interpretations of the results[33].

4. Empirical Strategy

I use a two-way error components model with individual-specific time-invariant unobserved heterogeneity (c_i) and survey wave indicators (d_t) for waves $t \in \{2003, \ldots, 2011\}$. That is,

$$\mathrm{Y_{it}} = \mathrm{W_i}\gamma + \mathrm{X_{it}}\beta + \mathrm{c_i} + d_t + u_{it} \quad (1)$$

where $\mathrm{Y_{it}}$ is an indicator equal to one if young adult i co-resides with their parent(s) at time *t*, $\mathrm{X}'_{\mathrm{it}}$ is a vector of time-varying factors, W'_i is a vector of time-invariant measurements, and $\mathrm{v_{it}} = \mathrm{c_i} + \mathrm{d_t} + u_{it}$ is not decomposed in the pooled regression setting. When decomposed, d_t is a collection of survey year dummies (each equal to one when t equals the corresponding wave's year) acknowledging year-specific unobserved heterogeneity common to cross-sectional units. Meanwhile, c_i is a vector of individual-specific, time-invariant, unobserved heterogeneity. Indeed, I suspect both time-variant and time-constant unobserved heterogeneity[34]. Lastly, u_{it}

[33] The research question may alter to: what factors discourage or encourage the decision to move out of the nuclear family? Nevertheless, I proceed with analysis as originally intended.

[34] For instance, macroeconomic events (such as housing market collapse or Great Recession exposure) are time-variant effects presumably common to all individuals, while child-rearing outcomes can be specific to the actions of an individual's parents and can incur long-term, time-constant effects. Child-rearing outcomes have implications on

denotes idiosyncratic disturbances particular to individual i at time t. A binary indicator as an outcome calls for the use of binary response models. However, it is known that linear probability models (LPMs) have inherent shortcomings, so the nonlinear binary choice models (Probit) will be considered as well. Pooled, random effects (RE), fixed effects (FE), and maximum likelihood estimation (MLE) are necessary.

Cross-sectional units of analysis are young adults (older than 17) defined to be members G1, each of which correspond to their biological mother and father in G0. Thus, X_{it} and W_i will need to contain variables pertaining to the G1 young adult, the G0 father, and the G0 mother, including relational measures among the three family members. The PSID is well-suited[35] for this generational analysis because family members are followed regardless of whether they live in the same household[36].

As previously mentioned, LPMs are known for their limitations[37] as well as for their strength in parsimonious interpretation as hedonic probability models. Despite the limitations that make LPMs less than ideal, LPMs serve as good baseline models when the outcome is a binary response. I preliminarily estimate the regression equation above under the LPM

---

unobservables such as diligence, role model preference, or industriousness (economic autonomy). Erik Erikson's famous stages of psychosocial development come to mind in this context.

[35] Family members are easily linked through an auxiliary component of the PSID called the Family Identification Mapping System (FIMS). That is, FIMS links G1 members to their G0 mothers and fathers. FIMS also allows us to follow both adoptive and biological parents of G1 members, but I can omit those with adoptive parents as this subsample only composes about two percent of G1 members in the PSID. Furthermore, FIMS also allows us to observe up to four generations contemporaneously (e.g., G3 members, G2 parents, G1 great-grandparents, G0 great-great grandparents can form one individual-year observation). Here, I only observe two generations as the same time (G1 members, G0 parents).

[36] The PSID distinguishes between family units and households. Households can contain multiple family units, thus multiple heads of family units. If the G1 individual moves out of the nuclear family unit (and not merely institutionalized such as the case of going off to college) *and* chooses to continue participating in the PSID, then the G1 member is assigned a new family unit identifier. If co-residence occurs at some time after, then the G1 member will continue being recognized as an autonomous family unit by the PSID. Thus, I exploit the methodology of the PSID for the purpose of this study.

[37] Inherent shortcomings to linear probability models include instances of predicted probabilities landing outside of the unit interval as well as an inherent constant partial effect assumption.

framework. Due to the nature of the outcome variable, however, partial effects that are contingent on regressor values (such as with nonlinear binary response models) are sensible; the inherent assumption of non-diminishing marginal returns under LPMs can be dubious depending on the relationship in question[38]. Proceeding with panel data estimation, I identify the significance of time-invariant unobserved heterogeneity by comparing pooled OLS (POLS) and random effects (RE) estimation results. Finally, I assess the necessity of fixed effects (FE) estimation by conducting the Mundlak-Hausman test using heteroskedasticity-robust RE point estimates and time averages of time-variant variables as additional regressors (Mundlak, 1978). This is equivalent to measuring whether the regressors in X_{it} are significantly correlated with c_i; if true, then RE estimation produces inconsistent (or asymptotically biased) estimates.

Nonlinear binary response estimation is underneath the *latent variable* framework:

$$\begin{cases} \mathrm{Y}_{\mathrm{it}}^{*} = \mathrm{W}_{\mathrm{i}}\gamma + \mathrm{X}_{\mathrm{it}}\beta + \mathrm{c}_{\mathrm{i}} + d_t + e_{it} \ ; e_{it} \overset{\text{i.i.d.}}{\sim} N(0,1) \\ \mathbf{Pr}[\mathrm{Y}_{\mathrm{it}} = 1 | (\mathrm{W}_{\mathrm{i}}, \mathrm{X}_{\mathrm{it}})] = \mathbf{\Phi}[\mathrm{W}_{\mathrm{i}}\gamma + X_{it}\beta + c_i + d_t] \end{cases} \tag{2}$$

where $\mathbf{\Phi}[*]$ is the standard normal distribution's CDF. While this setup demonstrates more mathematical obedience than LPMs and allows for the modeling of diminishing partial effects, nonlinear binary response models also have a drawback: estimation depends on imposing a strong distributional assumption on the unobserved disturbances[39]. Further, partial effects interpretation for panel Probit models requires us to relax to presence of time-invariant unobserved heterogeneity[40]. RE Probit estimation follows pooled Probit estimation, and a

[38] An illustration can be seen in the estimation of the effect of G0's maximum age on co-residence probability; it would be contestable that the effect of aging from 44 to 46 years would be associated with the same magnitude of an effect on co-residence probability as aging from 74 to 76 years. This is a reason for allowing for value-contingent partial effects, despite its trade-off with parsimonious interpretation.
[39] The idiosyncratic disturbances are assumed to be independent and identically standard normal-distributed.
[40] That is, we set $c_i = 0$ when computing either average partial effects or partial effects at the average. In my case, I compute the partial effects at the average (PEA).

comparison of estimates may reveal the significance of c_i. Note that one does not need to impose the assumption of zero correlation between the regressors and the unobserved heterogeneity; I can estimate the model again along with the time averages of X_{it} variables in an estimation method otherwise known as Chamberlain's RE Probit, which can be roughly described as a "quasi-fixed-effects" estimation using a control function approach[41]. I can then compare the Chamberlain estimates to the RE Probit estimates and conduct goodness-of-fit tests[42] to determine the best model. Rather than opting for a nonlinear FE model, I use the Chamberlain RE Probit's advantage in its ability to test c_i's statistical significance[43].

5. Results

After RE estimation, I model the composite residuals ($\widehat{v_{it}}$) as an AR(1) process under the strict exogeneity assumption[44] to ascertain whether RE estimation is indeed preferred over POLS. I find that autocorrelation among the composite errors is significant ($\hat{t} = 59.60$), which validates RE estimation. Realizing that the unobserved heterogeneity is statistically significant, I proceed further with FE estimation after validating the assumption of zero correlation between the observed and unobserved heterogeneity (under the RE framework) by conducting Mundlak's

[41] The time-constant unobserved heterogeneity is modeled as a function of the time averages of the regressors. That is, $c_i = \omega + \overline{X_i}\alpha + \varepsilon_i$, where $\frac{\varepsilon_i}{\sigma_u}|X_{it} \sim N(0,1)$. Chamberlain's RE Probit is comparable to the Mundlak (1978) Hausman test for RE strict exogeneity (regarding FE estimation validation). Chamberlain's Probit is justified if α is statistically different from $\vec{0}$.

[42] In the context of nonlinear binary response models, goodness-of-fit tests consist of computing the percentage of correctly predicted 0s and 1s, or even the overall percentage correctly predicted. A challenge to this process is determining the optimal threshold between success and failure probabilities. I use average predicted probability.

[43] While the FE Logit estimation entails a transformation that eliminates c_i, the significance of the fixed effect would not be testable. Testing the joint significance of the time averages of the regressors is equivalent to testing for c_i's significance, so I favor Chamberlain's Probit. What's more is that the Chamberlain Probit acts a "quasi-FE" estimation (similar to a within transformation, but not quite).

[44] That is, I test $H0\colon \rho = 0$ where $\widehat{v_{it}} = \rho\widehat{v_{i,t-1}} + error_{it}$. Under the null hypothesis, c_i is not significant, thus POLS is preferred. If c_i were significant, then autocorrelation among the composite error terms would exist. I would therefore prefer the decomposition of v_{it} under RE estimation (over POLS).

regression-based Hausman test[45]. I find that this assumption is invalid: the time averages of time-varying regressors are jointly significant ($\widehat{\chi^2_8} = 148.65$), rendering the RE estimator as likely inconsistent. After controlling and eliminating the unobserved heterogeneity, the effect of the mother's marital status increases in magnitude and significance. According to the FE estimate, a young adult with a married mother is *less* likely to co-reside with parents by 12.1 percentage points with other factors held constant. In other words, co-residence is significantly likelier if the mother is unmarried. This finding agrees with Pilkauskas & Cross (2018). Furthermore (with the FE estimate), Medicare coverage of at least one of a young adult's parents increases co-residence probability by 4.96 percentage points, ceteris paribus. Although this result remains significant when using robust (to heteroskedasticity and autocorrelation) standard errors, its sign changes once c_i is eliminated. The root of this sign reversal has yet to be determined. Finally, the goodness-of-fit metrics (overall percentage correctly predicted) for the linear models are quite low, suggesting that linear estimation may not be preferred to nonlinear estimation; I may need to allow for non-constant marginal effects.

In the nonlinear models (Table 2B), the mother's marital status and whether either parent is covered by Medicare both remain as statistically significant regressors. However, the co-residence effect of G0's Medicare coverage decreases to the 0.1 significance level. The estimate reverses in impact (from negative to positive) after modeling c_i as a function of $\overline{X}_\iota$, mirroring the same occurrence in the linear models[46]. According to the Chamberlain estimates, co-residence

---

[45] I test $H0\colon \alpha = \vec{0}$ where $y_{it} = W_i\gamma + X_{it}\beta + \overline{X}_\iota\alpha + c_i + d_t + u_{it}$ is estimated using RE. That is, α is the coefficient vector for the time averages of the time-varying regressors, $\overline{X}_\iota$. Under this null hypothesis, the RE estimator is preferred over the FE estimator. If the time averages are jointly significant, I prefer FE to RE.

[46] Another consequence of the Chamberlain Probit estimation is that parental homeownership loses significance. Furthermore, the impact of being covered by employer-provided health insurance is not a robust result, as seen from the HAC estimate. The sign reversal of G0's Medicare coverage needs an explanation or alleviation in future analysis.

probability for the average G1 adult is 9.21 percentage points higher if their mother is unmarried, and 3.88 percentage points higher if either G0 parent is covered by Medicare. Note that the Chamberlain RE Probit model yields the best performance with 78.52 overall percentage correctly predicted. Thus, in addition to allowing for diminishing marginal effects, modeling the time-constant unobserved heterogeneity as a function of the time averages of regressors seems to be a necessary approach in modeling co-residence probability underneath the RE framework[47].

6. Conclusion and Future Analysis

The marital status of a young adult's mother is significantly correlated with co-residence probability. I interpret aging, unmarried mothers to be economically vulnerable to adverse shocks; the negative association of her being married with co-residence probability is therefore sensible. Meanwhile, the positive association of G0's Medicare coverage with co-residence probability likely reflects its concurrence with the retirement decision (hence, G0's age). As of now, no causal claim can be made about these variables, though some insight has been shed.

Immediate future analysis should address some aspects of the data[48]. Relevant measurements such as immigrant status (or citizenship status) have not been identified within the PSID. Other variables that have yet to be derived from the PSID are birthplace (or where an individual grew up), whether one grew up in a single-parent household, and the number of siblings belonging to an individual. There are also some PSID measures that may be more

[47] The χ^2_8 test statistics are 187.50 and 159.31 when testing the joint significance of the time averages in the Chamberlain and HAC Chamberlain RE Probits, respectively.

[48] Technically speaking, some observations are dropped from estimation in the end. Specifically, 56 observations are dropped during estimation, including six entire cross-sectional units. Hence, there are 26 missing observations that render the panel unbalanced. Dropped observations are most likely due to missing values for some regressors. I therefore need to validate my data set for missing values. Some analysis regarding multicollinearity (such as analysis of variance inflation factors) is necessary, too.

insightful than others. For instance, it is better to use the PSID's home value or equity measure instead of a plain homeownership indicator to better assess whether generationally independent living arrangements are indeed a normal good. Plus, the level of attachment to the housing market (beyond one's status as a homeowner) remains unknown in this rendition of the study. Knowledge of housing market attachment could identify agents' heterogeneous responses to the 2007 housing market collapse or even exposure to the Great Recession. Another negligence is that the effect of the Affordable Care Act (which took effect on March 23, 2010) may play a significant role in health care measurements for the last survey wave. Finally, the definitions of some generational variables pertaining to G0 are negligent of differences between G0 mothers and fathers[49]. Better definitions for some variables in this study need to be determined.

There are also other estimation techniques to consider. For one, there is some cross-sectional correlation when siblings share parents; it may be necessary to account for this. Another model to consider is the linear Hausman-Taylor RE/FE mixture which preserves some important time-constant characteristics from the within transformation. However, I find from the present analysis that nonlinear modeling yields better results, indicating that it is necessary to allow for diminishing marginal effects on co-residence probability. Nevertheless, the conclusions drawn from the Chamberlain RE Probit model validate an FE approach, though in the nonlinear setting.

Although this paper is ultimately a study of correlation, a knowledge of variables that significantly move with or against co-residence outcomes may earn researchers some insight into potential underlying causal mechanisms for co-residence. Variables such as G0's Medicare

[49] For instance, the indicator for G0's Medicare coverage is equal to 1 if either G0 parent is covered by Medicare. However, whether the mother or the father (or both) are covered by Medicare is disregarded despite the potential significance of one parent being covered rather than the other.

coverage and whether G1's mother is married were found to be significantly correlated with co-residence probability. These correlates remained significant in robust linear (FE) and nonlinear (Chamberlain RE Probit) models, despite the sign reversal of Medicare coverage's effect. At this stage, I can at least confirm meaningful correlation between co-residence and factors mentioned by Pilkauskas & Cross (2018) and Pilkauskas et al. (2020), all from the perspective of the "squeezed middle" generation.

TABLE 1A: Summary statistics for time-constant variables

	Mean	SD	Min	Median	Max
White	.771	0.420	0	1	1
Black	.182	0.386	0	0	1
Female	.488	0.500	0	0	1
Mother graduated high school	.782	0.413	0	1	1
Father graduated high school	.783	0.412	0	1	1
Mother attained some college only	.18	0.384	0	0	1
Father attained some college only	.177	0.382	0	0	1
Mother completed 4 years of college	.183	0.387	0	0	1
Father completed 4 years of college	.277	0.447	0	0	1
Father’s education level	12.295	3.888	0	12	17
Mother’s education level	11.478	4.305	0	12	17

Unweighted summary statistics. N = 933, T = 5. Cross-sectional unit of analysis is an individual who is at least 18 years-old, is observed in all five PSID waves over 2003-2011, has biological parents only (who are observed in the PSID), and has no missing values for the variables of interest.

TABLE 1B: Summary statistics for time-varying variables

	Mean	SD	Min	Median	Max
Covered by employer-provided health insurance	.733	0.442	0	1	1
Parent(s) covered by employer-provided health insurance	.608	0.488	0	1	1
Father covered by Medicare	.428	0.495	0	0	1
Mother covered by Medicare	.338	0.473	0	0	1
Parent(s) covered by Medicare	.451	0.498	0	0	1
Age	36.564	10.385	18	36	66
Father’s age	64.517	11.410	0	64	97
Mother’s age	59.771	15.522	0	60	94
Education level	13.795	2.047	6	14	17
Employed	.803	0.398	0	1	1
On the job search	.064	0.245	0	0	1
Family head	.558	0.497	0	1	1
Spouse of the family head	.264	0.441	0	0	1
Child of the family head	.154	0.361	0	0	1
Real child care expenditures in 1000s of 2011 US dollars (G0)	.394	1.291	0	0	9.779
Had positive amount of childcare expenses	.141	0.348	0	0	1
Has 1 child in family unit	.207	0.405	0	0	1
Parents own a home	.887	0.317	0	1	1
Parents have a non-SS pension	.366	0.482	0	0	1
Parents’ real non-SS transfer income in 1000s of 2011 US dollars	6.817	10.290	0	0	48.896
Mother is married	.864	0.342	0	1	1
Co-reside	.196	0.397	0	0	1
Split-off	.052	0.222	0	0	1

Unweighted summary statistics. N = 933, T = 5. Cross-sectional unit of analysis is an individual who is at least 18 years-old, is observed in all five PSID waves over 2003-2011, has biological parents only (who are observed in the PSID), and has no missing values for the variables of interest. Interpretation of pooled means: average value in a given year (ensemble average of time averages). Expenditure/income variables deflated using percentage change in CPI with 2011 as base year.

TABLE 1C: Transition probabilities

Split-off		0	1	Total
	0	95.49	4.51	100
	1	97.81	2.19	100
Co-reside		0	1	Total
	0	96.95	3.05	100
	1	28.46	71.04	100
Has 1 child in family unit		0	1	Total
	0	90.30	9.70	100
	1	39.24	60.76	100
Has 2 children in family unit		0	1	Total
	0	92.91	7.09	100
	1	32.08	67.92	100
Parents have a non-SS pension		0	1	Total
	0	91.46	8.54	100
	1	20.77	79.23	100
Mother is married		0	1	Total
	0	93.81	6.19	100
	1	4.07	95.93	100
On the job search		0	1	Total
	0	95.12	4.88	100
	1	70.22	29.78	100
Parent(s) covered by Medicare		0	1	Total
	0	85.27	14.73	100
	1	6.39	93.61	100
Covered by employer-provided health insurance		0	1	Total
	0	74.07	25.93	100
	1	10.33	89.67	100
Parents own a home		0	1	Total
	0	86.87	13.13	100
	1	2.97	97.03	100

TABLE 2A: Linear Binary Response Model Results				
Y = 1[coreside]	**(1)** **POLS**	**(2)** **RE**	**(3)** **FE**	**(4)** **HAC FE**
Mother is married	0.0008	-0.0570**	-0.121***	-0.121***
	(0.0163)	(0.0181)	(0.0213)	(0.0273)
Parent(s) covered by Medicare	-0.148***	-0.0308*	0.0496***	0.0496***
	(0.0120)	(0.0127)	(0.0147)	(0.0144)
Has 1 child in family unit	-0.0009	0.0101	0.0108	0.0108
	(0.0136)	(0.0117)	(0.0122)	(0.0159)
Parents own a home	-0.115***	-0.0853***	-0.0491*	-0.0491
	(0.0177)	(0.0196)	(0.0230)	(0.0299)
Covered by employer-provided health insurance	-0.1280***	-0.0701***	-0.0343*	-0.0343
	(0.0130)	(0.0124)	(0.0136)	(0.0189)
State unemployment rate	-0.0006	-0.0089*	-0.0125**	-0.0125*
	(0.0042)	(0.0042)	(0.0046)	(0.0058)
Parents' non-SS transfer income (1000s 2011 US dollars)	-0.0008	-0.0019**	-0.0014*	-0.0014
	(0.0006)	(0.0006)	(0.0006)	(0.0007)
On the job search	0.1570***	0.0615***	0.0334	0.0334
	(0.0230)	(0.0181)	(0.0184)	(0.0267)
Constant	0.533***	0.537***	0.527***	0.527***
	(0.0374)	(0.0379)	(0.0409)	(0.0548)
Time Dummies	-	X***	X***	X***
N	927	927	927	927
Overall R^2	0.1051	0.082	0.0109	0.0109
Test stat for overall significance	67.52 $(F_{8,4600})$	320.86 (χ^2_{12})	24.02 $(F_{12,3670})$	11.31 $(F_{12,926})$
Fraction of $Var(c_i + u_{it})$ due to c_i	-	0.5569	0.6491	0.6491
Overall percentage correctly predicted	58.11%	60.11%	56.36%	56.36%

T = 5. HAC denotes estimation with a "heteroskedasticity and autocorrelation consistent" variance-covariance matrix. *, **, and *** denote significance of the effect at the 0.10, 0.05, and 0.01 levels, respectively. Time dummies use 2003 as the base year. Threshold value for all predictions is average predicted probability (per model).

TABLE 2B: Probit Binary Response Model Results: Partial Effects at the Average				
	(5)	**(6)**	**(7)**	**(8)**
Y = Pr[coreside]	**pooled**	**RE**	**Chamberlain**	**HAC Chamberlain**
Mother is married	0.0004	-0.0455**	-0.0921***	-0.0921***
	(0.0175)	(0.0166)	(0.0206)	(0.0267)
Parent(s) covered by Medicare	-0.1580***	-0.0535***	0.0388*	0.0388*
	(0.0125)	(0.0138)	(0.0154)	(0.0162)
Has 1 child in family unit	-0.00577	0.0124	0.0125	0.0125
	(0.0139)	(0.0109)	(0.0114)	(0.0128)
Parents own a home	-0.0991***	-0.0766***	-0.0354	-0.0354
	(0.0164)	(0.0176)	(0.0204)	(0.0243)
Covered by employer-provided health insurance	-0.115***	-0.0546***	-0.0250*	-0.0250
	(0.0122)	(0.0111)	(0.0114)	(0.0130)
State unemployment rate	0.0002	-0.0041	-0.0067	-0.0067
	(0.0044)	(0.0041)	(0.0045)	(0.0055)
Parents' non-SS transfer income (1000s 2011 US dollars)	-0.0010	-0.0016**	-0.0008	-0.0008
	(0.0006)	(0.0006)	(0.0006)	(0.0007)
On the job search	0.125***	0.0443**	0.0249	0.0249
	(0.0204)	(0.0143)	(0.0141)	(0.0160)
Time Dummies	X***	X***	X***	X***
N	927	927	927	927
Log-likelihood	-2006.67	-1399.84	-1314.68	-1314.68
LR Test Stat for Overall Significance	535.45 (χ^2_{12})	253.49 (χ^2_{12})	312.94 (χ^2_{20})	219.62 (χ^2_{20})
Fraction of $Var(c_i + u_{it})$ due to c_i	-	0.8445	0.831	0.831
Overall percentage correctly predicted	66.43%	73.72%	78.52%	78.52%

T = 5. HAC denotes estimation with a "heteroskedasticity and autocorrelation consistent" variance-covariance matrix. Displayed in the cells are the partial effects at the average (not the MLE point estimates). See Table 1B for averages. In the parentheses are the standard errors for the MLE point estimates. *, **, and *** denote significance of the effect at the 0.10, 0.05, and 0.01 levels, respectively. An intercept term was included when estimating each model. Threshold value for all predictions is average predicted probability (per model). Time dummies use 2003 as base year. For the Chamberlain models, the time averages of the regressors are jointly significant at the 0.01 level.

References

Altonji, J. G., Hayashi, F., & Kotlikoff, L. J. (1989). *Is the extended family altruistically linked? Direct tests using micro data*.

Bartczak, A., Budziński, W., Chilton, S., McDonald, R., & Nielsen, J. S. (2021). Altruism and efficient allocations in three-generation households. *Journal of Risk and Uncertainty*, *62*(2), 113–135.

Bastian, J., & Michelmore, K. (2018). The long-term impact of the earned income tax credit on children's education and employment outcomes. *Journal of Labor Economics*, *36*(4), 1127–1163.

Bengtson, V. L. (2001). Beyond the nuclear family: the increasing importance of multigenerational bonds: the burgess award lecture. *Journal of Marriage and Family*, *63*(1), 1–16.

Brown, J. R., & Finkelstein, A. (2011). Insuring long-term care in the United States. *Journal of Economic Perspectives*, *25*(4), 119–142.

de Bresser, J., Knoef, M., & Kools, L. (2021). Cutting one's coat according to one's cloth--How did the great recession affect retirement resources and expenditure goals? *Journal of Economic Behavior \& Organization*, *188*, 126–166.

Diamond, P. A. (1965). National debt in a neoclassical growth model. *The American Economic Review*, *55*(5), 1126–1150.

Engelhardt, G. V. (2008). Social security and elderly homeownership. *Journal of Urban Economics*, *63*(1), 280–305.

Engelhardt, G. V, Gruber, J., & Kumar, A. (2020). Early Social Security Claiming and Old-Age Poverty: Evidence from the Introduction of the Social Security Early Eligibility Age. *Journal of Human Resources*, 0119--9973R1.

Engelhardt, G. V, Gruber, J., & Perry, C. D. (2005). Social security and elderly living arrangements evidence from the social security notch. *Journal of Human Resources*, *40*(2), 354–372.

Gubernskaya, Z., & Tang, Z. (2017). Just like in their home country? A multinational perspective on living arrangements of older immigrants in the United States. *Demography*, *54*(5), 1973–1998.

Gustman, A. L., Steinmeier, T. L., & Tabatabai, N. (2010). What the stock market decline means for the financial security and retirement choices of the near-retirement population. *Journal of Economic Perspectives*, *24*(1), 161–182.

Harvey, H., Dunifon, R., & Pilkauskas, N. (2021). Under Whose Roof? Understanding the Living Arrangements of Children in Doubled-Up Households. *Demography*, *58*(3), 821–846.

Hayashi, F., Altonji, J., & Kotlikoff, L. (1996). Risk-sharing between and within families. *Econometrica: Journal of the Econometric Society*, 261–294.

Kahn, J. R., Goldscheider, F., & Garc\'\ia-Manglano, J. (2013). Growing parental economic power in parent--adult child households: Coresidence and financial dependency in the United States, 1960--2010. *Demography*, *50*(4), 1449–1475.

Kotlikoff, L. J., & Morris, J. N. (2008). *5. Why Don't the Elderly Live with Their Children? A New Look*. University of Chicago Press.

Lusardi, A., & Mitchell, O. S. (2007). Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics*, *54*(1), 205–224.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica: Journal of the Econometric Society*, 69–85.

Pilkauskas, N., & Michelmore, K. (2019). The effect of the Earned Income Tax Credit on housing and living arrangements. *Demography*, *56*(4), 1303–1326.

Pilkauskas, N. V, Amorim, M., & Dunifon, R. E. (2020). Historical Trends in Children Living in Multigenerational Households in the United States: 1870--2018. *Demography*, *57*(6), 2269–2296.

Pilkauskas, N. V, & Cross, C. (2018). Beyond the nuclear family: Trends in children living in shared households. *Demography*, *55*(6), 2283–2297.

Poterba, J. M. (2014). Retirement security in an aging population. *American Economic Review*, *104*(5), 1–30.

Rupert, P., & Zanella, G. (2018). Grandchildren and their grandparents' labor supply. *Journal of Public Economics*, *159*, 89–103.

Rust, J., & Phelan, C. (1997). How social security and medicare affect retirement behavior in a world of incomplete markets. *Econometrica: Journal of the Econometric Society*, 781–831.

Wiemers, E. E., Slanchev, V., McGarry, K., & Hotz, V. J. (2017). Living arrangements of mothers and their adult children over the life course. *Research on Aging*, *39*(1), 111–134.

Zhao, L., & Burge, G. (2017). Housing wealth, property taxes, and labor supply among the elderly. *Journal of Labor Economics*, *35*(1), 227–263.