Interdependent Durations in Joint Retirement*

Bo E. Honoré † Áureo de Paula ‡

Princeton University UCL, CeMMAP, IFS

This Version: February 2013

Abstract

In this paper we use a new duration model to study joint retirement in married couples using the Health and Retirement Study. Whereas conventionally used models cannot account for joint retirement, our model admits joint retirement with positive probability and nests the traditional proportional hazards model. In contrast to other statistical models for simultaneous durations, it is based on Nash bargaining and is interpretable as an economic behavior model. We provide a discussion of relevant identifying variation and estimate our model using indirect inference. *JEL Codes:* J26, C41, C3.

*We thank Richard Blundell, Martin Browning, Manuel Arellano, Liran Einav, Geert Ridder, Han Hong, Gerard van den Berg and Frank Wolak for useful conversations and David Card for suggesting this application. Financial support from a Steven H. Sandell grant from the Center for Retirement Research (Boston College) is gratefully acknowledged. Bo Honoré also acknowledges financial support from the National Science Foundation and from the Gregory C. Chow Econometric Research Program. Áureo de Paula also acknowledges financial support from the Economic and Social Research Council through the ESRC Centre for Microdata Methods and Practice grant RES-589-28-0001. Andrea Puig provided very able research assistance.

†Princeton University (Department of Economics), Princeton, NJ 08542. E-mail: honore@princeton.edu

‡University College London (Department of Economics), CeMMAP and IFS, London, UK. E-mail: apaula@ucl.ac.uk

1 Introduction and Related Literature

This paper investigates the determinants of joint retirement decisions in couples. A majority of retirees are married and many studies indicate that a significant proportion of individuals retires within a year of their spouse. Articles documenting joint retirement of couples (and datasets employed) include Hurd (1990) (New Beneficiary Survey); Blau (1998) (Retirement History Study); Gustman and Steinmeier (2000) (National Longitudinal Survey of Mature Women); Michaud (2003) and Gustman and Steinmeier (2004) (Health and Retirement Study); and Banks, Blundell, and Casanova Rivas (2007) (English Longitudinal Study of Ageing). Even though this is especially the case for couples closer in age, a spike in the distribution of retirement time differences at zero typically exists for most couples, regardless of the age difference. This is illustrated in Figure 1.

The spike in the distribution of the difference in the retirement dates for husbands and wives in Figure 1 suggests that many couples retire simultaneously. There are at least two distinct explanations for such a phenomenon. One is that husband and wife receive correlated shocks (observable or not), driving them to retirement at similar times. The other is that retirement is jointly decided, reflecting taste interactions of both members of the couple. In fact, 55% of respondents in the Health and Retirement Study expected to retire at the same time as their spouses.[1]

The distinction between these two drivers of joint retirement (which are not mutually exclusive) parallels the categorization by Manski (1993) of correlated and endogenous (direct) effects in social interactions. In that literature, the joint determination of a certain outcome of interest $y_i, i = 1, 2$ for two individuals $i = 1, 2$ is represented by the system of equations

$$\begin{aligned} y_1 &= \alpha y_2 + x_1^\top \beta + \epsilon_1 \\ y_2 &= \alpha y_1 + x_2^\top \beta + \epsilon_2 \end{aligned}$$

where x_i and $\epsilon_i, i = 1, 2$ represent observed and unobserved covariates determining y_i. We

[1]The figure corresponds to those who answer either YES or NO to the question: "Do you expect your spouse to retire at about the same time that you do?" (`R1RETSWP`). It excludes those whose spouse was not working.

want to separate the endogenous (direct) effect (α) from the correlation in ϵs. There, as in this article, discerning these two sources of correlation in outcomes is relevant for analytical and policy reasons. For example, if the estimated model does not allow for the joint decision by the couple, then the estimate of the effect of a retirement-inducing shock will be biased if the retirement times are indeed chosen jointly. Such spillover effects invalidate, for instance, the commonly employed Stable Unit Treatment Value Assumption (SUTVA) taken in the treatment effects literature, preventing the clear separation of direct and indirect effects occurring through feedback to the partner's retirement decision [e.g., Burtless (1990)]. Furthermore, the multiplier effect induced by the endogenous, direct effect of husband on wife or vice-versa is an important conduit for policy. The quantification of its relative importance is hence paramount for both methodological and substantive reasons.

Unfortunately, standard econometric duration models are not suitable to analyze joint durations with simultaneity of the kind that we have in mind. One tempting estimation strategy is to include the spouse's retirement date or, in the case of a hazard model, a time-varying variable indicating his or her retirement date. Because that is a choice which is potentially correlated with the unobservable variables determining a person's own retirement, estimators are bound to be inconsistent. Essentially this would amount to including an endogenous variable from a simultaneous equation model in the right hand side of a regression. An important contribution of this paper is therefore the specification of an econometric duration model that allows for simultaneity [see also de Paula (2009) and Honoré and de Paula (2010)]. As in the linear simultaneous equation model, identification is obtained using exclusion restrictions and, in our particular case, using the timing patterns in the data.

The broader literature on retirement is abundant and a number of papers focusing on retirement decisions in a multi-person household have appeared in the last 20 years. Hurd (1990) presents one of the early documentations of the joint retirement phenomenon. Later papers confirming the phenomenon and further characterizing the correlates of joint retirement are Blau (1998); Michaud (2003); Coile (2004a); Banks, Blundell, and Casanova Rivas (2007). Gustman and Steinmeier (2000) and Gustman and Steinmeier (2004) work with a dynamic economic model where husband and wife's preferences are affected by their spouses actions but make retirement decisions individually[2] and focus on Nash equilibria to the joint retirement decision, i.e. each spouse's retirement decision is optimal given the other spouse's timing and vice-versa.[3] More recently, Gustman and Steinmeier (2009) present a richer (non-unitary) economic model with a solution concept that differs from Nash Equilibrium and is guaranteed to exist and be unique. Michaud and Vermeulen (2011) estimate a version of the "collective" model introduced by Chiappori (1992) where (static) labor force participation decisions by husband and wife are repeatedly observed from a panel (i.e., the Health and Retirement Study). Casanova Rivas (2010) recently suggests a detailed unitary economic dynamic model of joint retirement. Coile (2004b) presents statistical evidence on health shocks and retirement decision by the couple and Blau and Gilleskie (2004) present an economic model also focusing on health outcomes and retirement in the couple.

In our analysis, we assume that retirement decisions are made through Nash Bargaining on the retirement date. This solution concept is attributed to Nash (1950) (though see also Zeuthen (1930)). It chooses retirement decisions to maximize the product of differences between spouses utilities and respective outside-options (i.e. "threat-points"). The Nash solution corresponds to a set of behavioral axioms on the bargaining outcomes (Pareto efficiency, independence of irrelevant alternatives and symmetry) and it is widely adopted in the literature on intra-household bargaining. It can be shown that this solution approximates the equilibrium outcome of a situation where husband and wife make offers to each other in an alternating order and the negotiation breaks down with a certain probability. As this probability goes to zero, the equilibrium converges to the Nash solution (see Binmore, Rubinstein, and Wolinsky (1986)). Though this solution also leads to Pareto efficient outcomes,

[2] In the Family Economics terminology, their model is a non-unitary model where people in the household make decisions individually. In unitary models, the household is viewed a single decision-making unit. A characterization of unitary and non-unitary models can found in Browning, Chiappori, and Lechene (2006).

[3] When more than one solution is possible, they select the Pareto dominant equilibrium, i.e. for all other equilibria at least one spouse would be worse off. In case no equilibrium is Pareto dominant, the equilibrium where retirement by at least one household member happens earliest is assumed (see, e.g., Gustman and Steinmeier (2000), pp. 515, 520).

it imposes more structure than Casanova Rivas (2010) or Michaud and Vermeulen (2011) [see Chiappori (1992) and Chiappori, Donni, and Komunjer (2012)].

Our model is a variation of a recently developed model (Honoré and de Paula (2010)) which extends well-known duration models to a (non-cooperative) strategic stopping game, where endogenous and correlated effects can be disentangled and interpreted (see also de Paula (2009) for a related analysis).[4] As such it is close to traditional duration models in the statistics and econometrics literature. Our model extends the usual statistical framework in a way that allows for joint termination of simultaneous spells with positive probability. In the usual hazard modeling tradition, this property does not arise. It is nonetheless essential to model joint retirement behavior. One can appeal to existing statistical models (e.g., Marshall and Olkin (1967)) to address this issue as done by An, Christensen, and Gupta (2004) in the analysis of joint retirement in Denmark, but parameter estimates cannot be directly interpretable in terms of the decision process by the couple. The framework presented in this paper directly corresponds to an economic model of decision-making by husband and wife and consequently can be more easily interpreted in light of such model. To estimate our model, we resort to indirect inference (Smith (1993); Gourieroux, Monfort, and Renault (1993); and Gallant and Tauchen (1996)), using as auxiliary models standard duration models and ordered models, as suggested in Honoré and de Paula (2010) for a similar model. (For an earlier application of indirect inference in a duration context, see Magnac, Robin, and Visser (1995)).

The remainder of this paper proceeds as follows. Section 2 describes our model and the empirical strategy for its estimation. In Section 3 we briefly describe the data and subsequently discuss our results in Section 4. We conclude in Section 5.

[4]In fact, our model estimates are obtained using auxiliary models that are interpretable in terms of the (non-cooperative game-theoretic) model in Honoré and de Paula (2010), which does not impose Pareto efficiency and equilibrium uniqueness.

2 Model and Empirical Strategy

In our model, spouse i receives a utility flow of $K_i > 0$ before retirement. After retirement, the utility flow is given by $Z_i(s)\varphi_i\delta(s, t_j)e^{-\rho s}$ at time s. The function $Z_i(\cdot)$ is an increasing function such that $Z_i(0) = 0$. In principle, it is possible to allow for kinks or discontinuities in $Z_i(\cdot)$. In a model without interdependence, those would correspond to discontinuities in the hazard rate in the case of kinks in $Z_i(\cdot)$ or, in the case of discontinuities in $Z_i(\cdot)$, positive probability of retirement at the discontinuity date. The factor $\varphi_i = \varphi_i(x_i)$ is a positive function of individual observable covariates. Time is discounted at the rate $\rho > 0$ and $\delta(s, t_j) = (\delta - 1)\mathbf{1}(s \geq t_j) + 1$ where $\delta > 1$ and t_j is the retirement date for spouse j, representing the effect of spouse j's retirement on i's utility flow from retirement. This structure is similar to the one defined in Honoré and de Paula (2010): if $\delta = 1$, we obtain a standard proportional hazards model for the time until retirement. The δ could be made player-specific as well, but we focus on homogenous δ for simplicity. Time is measured in terms of "family age," which is set to zero when the oldest partner in the couple reaches 60 years-old. Given a realization (k_1, k_2) for the random vector (K_1, K_2), retirement timing is obtained as the solution to the Nash bargaining problem [Nash (1950), see also Zeuthen (1930)]:

$$\begin{aligned} \max_{t_1,t_2} \left(\int_0^{t_1} k_1 e^{-\rho s} ds \right. & \left. + \int_{t_1}^{\infty} Z_1(s)\varphi_1\delta(s \geq t_2)e^{-\rho s} ds - A_1 \right) \times \\ & \left(\int_0^{t_2} k_2 e^{-\rho s} ds + \int_{t_2}^{\infty} Z_2(s)\varphi_2\delta(s \geq t_1)e^{-\rho s} ds - A_2 \right) \end{aligned} \tag{1}$$

where A_1 and A_2 are the threat points for spouses 1 and 2, respectively. In our estimation, we set A_i equal to a multiple of the utility spouse i would obtain if spouse j never retired. In the general setting there may also be asymmetric bargaining weights which appear as exponents in the objective function. We ignore this for simplicity, though our analysis could be generalized to include that case. The Nash bargaining solution corresponds to a set of behavioral axioms on the bargaining outcomes (Pareto efficiency, independence of irrelevant alternatives and symmetry) and it is widely adopted in the literature on intra-household bargaining. It can be shown that this solution approximates the equilibrium outcome of a situation where husband and wife make offers to each other in an alternating order and the negotiation breaks down with a certain probability. As this probability goes to zero, the equilibrium converges to the Nash solution (see Binmore, Rubinstein, and Wolinsky (1986)). Note that the first term in (1) can be further simplified to

$$\begin{aligned} & \left(k_1 \left.\frac{-e^{-\rho s}}{\rho}\right|_0^{t_1} + \varphi_1 \int_{t_1}^{\infty} Z_1(s) e^{-\rho s} ds + \varphi_1 (\delta - 1) \int_{\max\{t_1,t_2\}}^{\infty} Z_1(s) e^{-\rho s} ds - A_1\right) \\ = & \left(k_1 \rho^{-1} \left(1 - e^{-\rho t_1}\right) + \varphi_1 \widetilde{Z}_1(t_1) + \varphi_1 (\delta - 1) \widetilde{Z}_1 (\max\{t_1, t_2\}) - A_1\right) \end{aligned}$$

where $\widetilde{Z}_i(t) = \int_t^{\infty} Z_i(s) e^{-\rho s} ds$ and hence $\widetilde{Z}_i'(t) = -Z_i(t) e^{-\rho t}$. An analogous simplification applies to the second term.

In the absence of an interaction effect ($\delta = 1$), a Weibull baseline hazard for the proportional hazard model would correspond to $Z_i(t) = Z(t; \alpha_i) = t^{\alpha_i}$ so in that case

$$\widetilde{Z}_i(t) = \int_t^{\infty} s^{\alpha_i} e^{-\rho s} ds = \left(\frac{1}{\rho}\right)^{\alpha_i + 1} \Gamma(\alpha_i + 1, \rho t)$$

where the upper incomplete gamma function is defined by $\Gamma(\alpha_i, x) = \int_x^{\infty} s^{\alpha_i - 1} e^{-s} ds$.[5]

In summary, the objective function is given by

$$\begin{aligned} N(t_1, t_2) = & \overbrace{\left(k_1 \rho^{-1} \left(1 - e^{-\rho t_1}\right) + \varphi_1 \widetilde{Z}_1(t_1) + \varphi_1 (\delta - 1) \widetilde{Z}_1(\max\{t_1, t_2\}) - A_1\right)}^{\equiv I} \times \\ & \underbrace{\left(k_2 \rho^{-1} \left(1 - e^{-\rho t_2}\right) + \varphi_2 \widetilde{Z}_2(t_2) + \varphi_2 (\delta - 1) \widetilde{Z}_2(\max\{t_1, t_2\}) - A_2\right)}_{\equiv II} \end{aligned}$$

If spouses retire sequentially, the objective function first order conditions are obtained as follows. Assuming $t_1 < t_2$ and taking derivatives with respect to t_1 we get:

$$\left(k_1 e^{-\rho t_1} - Z_1(t_1) \varphi_1 e^{-\rho t_1}\right) \left(\int_0^{t_2} k_2 e^{-\rho s} ds + \int_{t_2}^{\infty} Z_2(s) \varphi_2 \delta(s \geq t_1) e^{-\rho s} ds - A_2\right) = 0.$$

[5]This expression can be further manipulated by noting that if the random variable X is Gamma distributed with parameters α and $\beta = 1$

$$\overline{F}_{\Gamma(\alpha,1)}(x) = P(X > x) = \frac{1}{\Gamma(\alpha)} \int_x^{\infty} s^{\alpha - 1} e^{-s} ds = \frac{\Gamma(\alpha, x)}{\Gamma(\alpha)}.$$

Consequently,

$$\widetilde{Z}(t; \alpha) = \left(\frac{1}{\rho}\right)^{\alpha + 1} \Gamma(\alpha + 1, \rho t) = \left(\frac{1}{\rho}\right)^{\alpha + 1} \Gamma(\alpha + 1) \overline{F}_{\Gamma(\alpha + 1, 1)}(\rho t)$$

which is useful since both $\Gamma(\cdot)$ and $\overline{F}_{\Gamma(\cdot,1)}(\cdot)$ are preprogrammed in many software packages.

This implies that

$$k_1 = Z_1(t_1)\varphi_1$$

or

$$\int_0^{t_2} k_2 e^{-\rho s} ds + \int_{t_2}^{\infty} Z_2(s)\varphi_2 \delta(s \geq t_1) e^{-\rho s} ds = A_2.$$

The second possibility is ruled out since player 2 should get more than his or her threat point at an interior optimum. The first order condition with respect to t_2 gives

$$Z_2(t_2)e^{-\rho t_2}\varphi_1 (1-\delta) \times (II) + (I) \times \left(k_2 e^{-\rho t_2} - Z_2(t_2)\varphi_2 \delta e^{-\rho t_2}\right) = 0. \quad (2)$$

We note that the t_2 that sets the above expression to zero occurs earlier than the value obtained in Honoré and de Paula (2010): $Z_2^{-1}(k_2/\varphi_2\delta)$.[6] Because $Z_1(t_2)e^{-\rho t_2}\varphi_1 (1-\delta) \times (II) \leq 0$ at the optimum, for the first order condition to be zero the product $(I) \times (k_2 e^{-\rho t_2} - Z_2(t_2)\varphi_2 \delta e^{-\rho t_2})$ should be positive. If the product were zero, one would have $t_2 = Z_2^{-1}(k_2/\varphi_2\delta)$ (since setting I to zero would not be optimal as in previous arguments and we then have that $(k_2 e^{-\rho t_2} - Z_2(t_2)\varphi_2\delta e^{-\rho t_2}) = 0$, which is equivalent to $t_2 = Z_2^{-1}(k_2/\varphi_2\delta)$). To make the product positive, we then have to lower t_2 below $Z_2^{-1}(k_2/\varphi_2\delta)$. This implies that

$$\begin{aligned} T_1 &= Z_1^{-1}(K_1/\varphi_1) \\ T_2 &\leq Z_2^{-1}(K_2/(\varphi_2\delta)) \end{aligned}$$

which gives the same timing choice for the first retiree as in Honoré and de Paula (2010) but an earlier one for the second spouse. A similar set of calculations is obtained for $T_2 < T_1$.[7]

[6]In Honoré and de Paula (2010), we also require that $Z_1(\cdot) = Z_2(\cdot)$.

[7]For computation purposes we also notice that the objective function is unimodal on t_2. If we start at the critical value, increasing t_2 reduces the function. This is because, for small ρ, $Z_1(t_2)e^{-\rho t_2}\varphi_1 (1-\delta)$ becomes *more* negative and II becomes *more* positive, so the product becomes *more* negative. For the second term, I decreases and $k_2 e^{-\rho t_2} - Z_2(t_2)\varphi_2\delta e^{-\rho t_2}$, *which is positive*, decreases. Their product then decreases. Consequently, the derivative, which is the sum of these two products becomes negative, and the objective function is decreasing. Analogously we can also determine that the objective function is increasing for values below the critical value.

A third possibility is for spouses to retire jointly. In this case,

$$
\begin{aligned}
T &= \arg\max_t N(t,t) \\
&= \arg\max_t \left(k_1\rho^{-1}\left(1-e^{-\rho t}\right)+\varphi_1\widetilde{Z}_1(t)+\varphi_1(\delta-1)\widetilde{Z}_1(t)-A_1\right) \\
&\quad \left(k_2\rho^{-1}\left(1-e^{-\rho t}\right)+\varphi_2\widetilde{Z}_2(t)+\varphi_2(\delta-1)\widetilde{Z}_2(t)-A_2\right) \\
&= \arg\max_t \left(k_1\rho^{-1}\left(1-e^{-\rho t}\right)+\varphi_1\delta\widetilde{Z}_1(t)-A_1\right)\left(k_2\rho^{-1}\left(1-e^{-\rho t}\right)+\varphi_2\delta\widetilde{Z}_2(t)-A_2\right).
\end{aligned}
$$

The derivative of this is

$$
\begin{aligned}
& e^{-\rho t}\left(K_1-\varphi_1\delta Z_1(t)\right)\left(k_2\rho^{-1}\left(1-e^{-\rho t}\right)+\varphi_2\delta\widetilde{Z}_2(t)-A_2\right) \\
& +e^{-\rho t}\left(k_1\rho^{-1}\left(1-e^{-\rho t}\right)+\varphi_1\delta\widetilde{Z}_1(t)-A_1\right)\left(k_2-\varphi_2\delta Z_2(t)\right)
\end{aligned}
$$

which, set to zero, delivers the optimum implicitly. It can be noted that when $t<Z_1^{-1}\left(k_1/\left(\varphi_1\delta\right)\right)$ and $t<Z_2^{-1}\left(k_2/\left(\varphi_2\delta\right)\right)$ this is positive, and when $t>Z_1^{-1}\left(k_1/\left(\varphi_1\delta\right)\right)$ and $t>Z_2^{-1}\left(k_2/\left(\varphi_2\delta\right)\right)$ it is negative. The optimum is therefore in the interval

$$
\min\left\{Z_1^{-1}\left(k_1/\left(\varphi_1\delta\right)\right),Z_2^{-1}\left(k_2/\left(\varphi_2\delta\right)\right)\right\}\le t\le\max\left\{Z_1^{-1}\left(k_1/\left(\varphi_1\delta\right)\right),Z_2^{-1}\left(k_2/\left(\varphi_2\delta\right)\right)\right\}
$$

This is useful in the numerical solution to the above equation used in the estimation.

Figure 2 illustrates these cases and plots both T_1 and T_2 as a function of K_2 as K_1 is held fixed. For low values of K_2, $T_1>T_2$: labor force attachment is higher for spouse 1 than for spouse 2. When K_1 is large, on the other hand, $T_1<T_2$ and spouse 1 retires sooner. In intermediary values of K_1, $T_1=T_2$ and the two spouses retire at the same time. This generates probability distributions such as those in Figure 3. Unconditionally, the probability density function for T_1 is smooth. Conditionally on $T_2=t_2$ though, a point mass at $T_1=t_2$ arises.

The set of realizations of (K_1,K_2) for which $T=T_1=T_2$ is an optimum is larger than the set obtained in the non-cooperative setup from Honoré and de Paula (2010). This is illustrated in Figure 4, where the area between the dotted lines is the joint retirement region in Honoré and de Paula (2010) and the area between solid lines is the joint retirement region in the current paper. Also, whereas in that paper any date within a range $[\underline{T}<\overline{T}]$ (where $\underline{T}=\max\left\{Z_1^{-1}\left(k_1/\left(\varphi_1\delta\right)\right),Z_2^{-1}\left(k_2/\left(\varphi_2\delta\right)\right)\right\}$)was sustained as an equilibrium for pairs

(k_1, k_2) inducing joint retirement, in the current article the equilibrium joint retirement date for a given realization of (K_1, K_2) is uniquely pinned down. Because Nash bargaining implies Pareto efficiency and because $\underline{T}$ is the Pareto dominant outcome among the possible multiple equilibria in the game analyzed by Honoré and de Paula (2010), it should be the case that joint retirement in the Nash bargaining model occurs on or before $\underline{T}$. In comparison to the non-cooperative paradigm adopted in our previous paper, Nash bargaining allows spouses to "negotiate" an earlier retirement date, which is advantageous to both.

Finally, we note that when $\delta = 1$ the optimal retirement dates will correspond to

$$\log Z_i(t_i) = -\log \varphi_i + \log K_i, \quad i = 1, 2.$$

When K_i follows a unit exponential distribution, this corresponds to a proportional hazard model. For a general distribution of K_i, this yields the Generalized Accelerated Failure Time model of Ridder (1990).

2.1 Discussion of Identifying Variation

In this subsection we discuss informally the variation in the data that allows us to (non-parametrically) identify the elements of the model. First, note that the functions $Z_i(\cdot)$ and $\varphi_i(\cdot)$ are identified (up to scale) if covariates have a support large enough so that

$$\varphi_j \equiv \varphi_j(x_j) \to 0$$

as x_j is driven to the boundary of the support (possibly infinity). This implies that it is optimal to have $t_j \to \infty$: retirement age is arbitrarily large for that spouse. Intuitively, this would come about if the explanatory variables take values that make one of the spouses strongly attached to the labor force given his or her covariate values. (In our data, for example, about 5% of the persons retire past 72 years (i.e. 874 months) of age, which is much later than the eligibility age for early retirement.) The other spouse will then optimally retire at T_i such that

$$\log Z_i(T_i) = -\log \varphi_i(x_i) + \log K_i$$

and one can apply the arguments in Ridder (1990) to identify $Z_i(\cdot)$, $\varphi_i(\cdot)$ and the marginal distribution of K_i (up to scale). We note also that this identification argument operates irrespective of the values of A_1 and A_2 (or asymmetries in the bargaining power).

Having identified $Z_i(\cdot), \varphi_i(\cdot)$ and the marginal distribution of K_i, the interaction parameter δ is pinned down by the probability of joint retirement. To see this, note that $\delta = 1$ implies that the objective function is:

$$N(t_1, t_2) = \overbrace{\left(k_1\rho^{-1}\left(1 - e^{-\rho t_1}\right) + \varphi_1 \widetilde{Z}_1(t_1) - A_1\right)}^{\equiv I} \times \underbrace{\left(k_2\rho^{-1}\left(1 - e^{-\rho t_2}\right) + \varphi_2 \widetilde{Z}_2(t_2) - A_2\right)}_{\equiv II}.$$

The first order conditions with respect to t_1 and t_2 are

$$e^{-\rho t_1}(k_1 - \varphi_1 Z_1(t_1)) \times II = 0 \;\;\Rightarrow\;\; t_1 = Z_1^{-1}(k_1/\varphi_1)$$
$$e^{-\rho t_2}(k_2 - \varphi_2 Z_2(t_2)) \times I = 0 \;\;\Rightarrow\;\; t_2 = Z_2^{-1}(k_2/\varphi_2).$$

Then joint retirement ($t_1 = t_2$) would imply $Z_1^{-1}(K_1/\varphi_1) = Z_2^{-1}(K_2/\varphi_2)$, which has zero probability if (K_1, K_2) is continuously distributed.

On the other hand, if $\delta > 1$, $Pr(T_1 = T_2|x_1, x_2) > 0$. This can be seen by remembering that $T_1 < T_2$ implies

$$\left.\begin{array}{rcl} T_1 & = & Z_1^{-1}(K_1/\varphi_1) \\ T_2 & \leq & Z_2^{-1}(K_2/(\delta\varphi_2)) \end{array}\right\} \Rightarrow Z_1^{-1}(K_1/\varphi_1) < Z_2^{-1}(K_2/(\delta\varphi_2)).$$

Likewise, if $T_1 > T_2$ we have that $Z_2^{-1}(K_2/\varphi_2) < Z_1^{-1}(K_1/(\delta\varphi_1))$. These two implications are equivalently written as

$$Z_1^{-1}(K_1/\varphi_1) \geq Z_2^{-2}(K_2/(\delta\varphi_2)) \;\;\Rightarrow\;\; T_1 \geq T_2$$
$$Z_2^{-1}(K_2/\varphi_2) \geq Z_1^{-1}(K_1/(\delta\varphi_1)) \;\;\Rightarrow\;\; T_2 \geq T_1,$$

which in turn implies that

$$0 < Pr(Z_2(Z_1^{-1}(K_1/(\delta\varphi_1))\varphi_2 \leq K_2 \leq Z_2(Z_1^{-1}(\varphi_1))\delta\varphi_2) \leq Pr(T_1 = T_2|x_1, x_2),$$

where the first inequality follows if $\delta > 1$.[8] Intuitively, larger values of δ will induce joint retirement more likely and joint retirement will be informative about δ. Similarly, even

[8] In this case, $K_1/(\delta\varphi_1) < K_1/\delta\varphi_1 \Rightarrow Z_2(Z_1^{-1}(K_1/(\delta\varphi_1)) < Z_2(Z_1^{-1}(K_1/\varphi_1) < Z_2(Z_1^{-1}(K_1/\varphi_1))\delta$.

in the event of sequential retirement, whereas the first spouse to retire always retires at $Z_i^{-1}(K_i/\varphi_i)$, larger values of δ will lead to earlier retirement by the second spouse to retire providing variation to identify δ (see Appendix for details).

Finally, if $\mathbf{x}_1 \neq \mathbf{x}_2$ and the support of covariates is large enough, we note that for any (bounded convex) set of pairs (k_1, k_2) there is a pair (φ_1, φ_2) that induces sequential retirement. When realizations of (K_1, K_2) induce sequential retirement, the retirement dates t_1 and t_2 are a one-to-one mapping from k_1 and k_2. If $t_1 < t_2$, for example, t_1 is a one-to-one mapping of k_1 (i.e., $t_1 = Z_1^{-1}(k_1/\varphi_1)$). Given k_1 and k_2 (and consequently $t_1 = Z_1^{-1}(k_1/\varphi_1)$), t_2 is uniquely determined (see footnote 4). From the FOC, it is also clear that, given (t_1, t_2) (and $k_1 = Z_1(t_1)\varphi_1$), one can uniquely retrieve the corresponding k_2. Using the Jacobian method, one can see that the joint density of (T_1, T_2) should be informative about the joint density of (K_1, K_2). Intuitively, a different distribution of (K_1, K_2) on that (bounded convex) set changes the probability of (T_1, T_2) (the image of that set) given the covariates corresponding to the initial choice of (φ_1, φ_2). (Because the Jacobian transformation in the mapping between the two joint densities does not factor, we should also mention that even when K_1 and K_2 are independent, T_1 and T_2 are not (locally) independent on the $T_1 \neq T_2$ region.)

An alternative to the Nash bargaining framework we present here would be an utilitarian aggregation of the utility functions in the household (i.e., the collective model of Chiappori (1992)). In that case, the retirement dates would solve:

$$\max_{t_1,t_2} cU^1(t_1, t_2; \mathbf{x}_1, K_1) + U^2(t_1, t_2; \mathbf{x}_2, K_2),$$

where c stands for the relative weight of agent 1's utility. This leads to the following first-order conditions:

$$c \times \frac{\partial U^1(t_1, t_2; \mathbf{x}_1, K_1))}{\partial t_i} + \frac{\partial U^2(t_1, t_2; \mathbf{x}_2, K_2)}{\partial t_i} = 0, \quad i = 1, 2.$$

The setting we propose focuses instead on maximizing $(U^1(t_1, t_2; \mathbf{x}_1, K_1) - A_1) \times (U^2(t_1, t_2; \mathbf{x}_2, K_2) - A_2)$. This leads to the following first-order conditions:

$$\frac{U^2(t_1, t_2; \mathbf{x}_2, K_2) - A_2}{U^1(t_1, t_2; \mathbf{x}_1, K_1) - A_1} \times \frac{\partial U^1(t_1, t_2; \mathbf{x}_1, K_1))}{\partial t_i} + \frac{\partial U^2(t_1, t_2; \mathbf{x}_2, K_2)}{\partial t_i} = 0, \quad i = 1, 2.$$

Consequently, the two are equilivalent only if

$$c = \frac{U^2(t_1, t_2; \mathbf{x}_2, K_2) - A_2}{U^1(t_1, t_2; \mathbf{x}_1, K_1) - A_1}.$$

In this sense, the Nash bargaining setup imposes further restrictions on the model as pointed out by Chiappori, Donni, and Komunjer (2012). That paper also establishes identification results when a common set of covariates $\overline{\mathbf{x}}$ affects both the threat points $A_i, i = 1, 2$ and utilities $U_i, i = 1, 2$. Point-identification is achieved using spouse-specific covariates that affect the threat points $A_i, i = 1, 2$ but are excluded from $U^i, i = 1, 2$. In our empirical investigation we rely instead on spouse specific covariates in $U^i, i = 1, 2$ and no excluded variables in the threat point functions $A_i, i = 1, 2$. Moreover, Chiappori, Donni, and Komunjer (2012) assume that latent variables (i.e., $K_i, i = 1, 2$) are additively separable, which is not our case.

2.2 Estimation: Indirect Inference

Because the likelihood for this model is not easily computed in closed form, we resort to simulation assisted methods. One potential strategy would be to use Simulated Maximum Likelihood, where one nonparametrically estimates the conditional likelihood via kernel methods applied to simulations of T_1 and T_2 at particular parameter values and searches for the parameter value that maximizes the (simulated) likelihood. We opt for a different strategy for two main reasons. First, our likelihood displays some non-standard features. For example, there is a positive probability for the event $\{T_1 = T_2\}$. Second, consistency of the SML estimator requires a large number of simulations, which can be computationally expensive.

To estimate our model we employ an indirect inference strategy (see Gourieroux, Monfort, and Renault (1993); Smith (1993); and Gallant and Tauchen (1996)). Rather than estimating the Maximum Likelihood Estimator for the true model characterized by parameter θ, one estimates an approximate (*auxiliary*) model with parameter β. Let $n = 1, \ldots, N$ index a sample of households (couples). Then, under the usual regularity conditions,

$$\widehat{\beta} = \arg\max_b \sum_{n=1}^{N} \log \mathcal{L}_a(b; z_n) \xrightarrow{p} \arg\max_b E_{\theta_0}\left[\log \mathcal{L}_a(b; z_n)\right] \equiv \beta_0(\theta_0) \tag{3}$$

where $\mathcal{L}_a$ is the likelihood function for the auxiliary model and the expectation E_{θ_0} is taken with respect to the true model. $\beta_0(\theta_0)$ is known as the pseudo–true value and the key is that it depends on the true parameters of the data–generation process (θ_0). If one knew the pseudo–true value as a function of θ_0, then it could be used to solve the equation

$$\widehat{\beta} = \beta_0\left(\widehat{\theta}\right)$$

and obtain an estimator for θ_0. In our case, we do not know $\beta_0(\theta)$, but we can easily approximate this function using simulations. For each θ, we generate R draws

$$\left\{\left(z_{1r}\left(\theta\right), z_{2r}\left(\theta\right), \ldots, z_{Nr}\left(\theta\right)\right)\right\}_{r=1}^{R}$$

and then estimate the function

$$\beta_0(\theta) \equiv \arg\max_b E_\theta\left[\log \mathcal{L}_a\left(b; z_n\right)\right]$$

by

$$\widetilde{\beta}_R\left(\theta\right) = \arg\max_b \frac{1}{R}\sum_{r=1}^{R}\frac{1}{N}\sum_{n=1}^{N}\left(\log \mathcal{L}_a\left(b; z_{nr}\left(\theta\right)\right)\right)$$

In other words, we find $\widehat{\theta}$ such that the generated data set using $\widehat{\theta}$ gives the same estimate in the auxiliary model as we got in the real sample:

$$\widehat{\beta} = \widetilde{\beta}_R\left(\widehat{\theta}\right)$$

One could then estimate θ by making the difference between $\widehat{\beta}$ and $\widetilde{\beta}_R\left(\theta\right)$ as small as possible.

Expression (3) implies that

$$\widehat{\beta} = \arg\max_b \sum_{n=1}^{N} \log \mathcal{L}_a\left(b; z_n\right) \quad \Longrightarrow \quad \frac{1}{N}\sum_{n=1}^{N} \mathcal{S}_a\left(\widehat{\beta}; z_n\right) = 0,$$

so $\widehat{\beta}$ converges to the solution to

$$E_\theta\left[\mathcal{S}_a\left(b; z_n\right)\right] = 0,$$

which is just $\beta_0\left(\theta_0\right)$ from equaton (3). So if we knew the function $\beta_0\left(\theta\right)$ we would estimate θ_0 by solving $\widehat{\beta} = \beta_0\left(\widehat{\theta}\right)$ which is the same as

$$E_{\widehat{\theta}}\left[\mathcal{S}_a\left(\widehat{\beta}; z_n\right)\right] = 0.$$

As before, we estimate $E_\theta\left[\mathcal{S}_a\left(\cdot; z_n\right)\right]$ as a function of θ using

$$\frac{1}{R}\sum_{r=1}^{R}\frac{1}{N}\sum_{n=1}^{N}\mathcal{S}_a\left(\cdot; z_{nr}\left(\theta\right)\right)$$

and θ_0 is estimated by solving

$$\frac{1}{R}\sum_{r=1}^{R}\frac{1}{N}\sum_{n=1}^{N}\mathcal{S}_a\left(\widehat{\beta}; z_{nr}\left(\theta\right)\right) = 0.$$

If $\dim\left(\mathcal{S}_a\right) > \dim\left(\beta\right)$, we minimize

$$\left(\frac{1}{R}\sum_{r=1}^{R}\frac{1}{N}\sum_{n=1}^{N}\mathcal{S}_a\left(\widehat{\beta}; z_{nr}\left(\theta\right)\right)\right)^{\top} W \left(\frac{1}{R}\sum_{r=1}^{R}\frac{1}{N}\sum_{n=1}^{N}\mathcal{S}_a\left(\widehat{\beta}; z_{nr}\left(\theta\right)\right)\right)$$

over θ. The weighting matrix W is a positive definite matrix performing the usual role in terms of estimator efficiency. This strategy is useful because we only estimate the auxiliary model once using the real data. After that, we evaluate its FOC for different draws of θ.

2.2.1 Auxiliary Model

Our auxiliary model is composed of three reduced form models that are chosen to capture the features of the data that are our main concern: the duration until retirement for each spouse, and the idea that members of some married couples choose to retire jointly. For the first two, we use a standard proportional hazard model for each spouse with a Weibull baseline hazard and the usual specification for the covariate function. For the third, we use an ordered Logit model as suggested by our paper Honoré and de Paula (2010). We present the models in detail below.

Weibull Proportional Hazard Model

For each spouse i, the hazard for retirement conditional on x is assumed to be $\lambda_i\left(t|\, x\right) = \alpha_i t^{\alpha_i - 1}\exp\left(x^{\top}\beta_i\right)$. The (log) density of retirement for spouse i conditional on x, $\log f_i\left(t|x\right)$, is then given by:

$$\log\left\{\lambda_i\left(t\right)\exp\left(x^{\top}\beta_i\right)\exp\left(-Z_i\left(t\right)\exp\left(x^{\top}\beta_i\right)\right)\right\} = \log\alpha_i + \left(\alpha_i - 1\right)\log t + x^{\top}\beta_i - t^{\alpha_i}\exp\left(x^{\top}\beta_i\right)$$

The (conditional) survivor function can be analogously obtained and is given by:

$$\log S_i\left(t|x\right) = \log\left\{\exp\left(-Z_i\left(t\right)\exp\left(x^{\top}\beta_i\right)\right)\right\} = -t^{\alpha_i}\exp\left(x^{\top}\beta_i\right)$$

In our dataset, about 55% of the individual retirement dates are censored (see Table 1). Letting $c_{i,n} = 1$ if the observed retirement date for spouse i in household n is (right-)censored, and $= 0$ otherwise, we obtain the log-likelihood function:

$$\log \mathcal{L}_i = \sum_{n=1}^{N} (1 - c_{i,n}) \left(\log \alpha_i + (\alpha_i - 1) \log (t_{i,n}) + x'_{i,n}\beta_i \right) - \sum_{n=1}^{N} t_{i,n}^{\alpha_i} \exp \left(x'_{i,n}\beta_i \right)$$

First and second derivatives used in the computation of the MLE for this auxiliary model are presented in the Appendix.

We note that, even though we assume that the censoring time is independent from retirement dates, we do not need to model the distribution of censoring times to simulate the model. When drawing observations from the model to fit the auxiliary duration models, we are able to censor the simulated observations using the date when respondents were last interviewed or died as censoring dates even for those who retire earlier than that in the data (i.e., the "censoring month" in Table 1).

Ordered Logit Model Pseudo MLE

In the spirit of the estimation strategy suggested in Honoré and de Paula (2010), we also use as auxiliary model an ordered logit. Whereas the Weibull model will convey information on the timing of retirement, this second auxiliary model will provide information on the pervasiveness of joint retirement and help identify the taste interactions leading to this phenomenon (i.e. δ). Define

$$y_n = \begin{cases} 1, & \text{if } t_1 > t_2 + 1 \\ 2, & \text{if } |t_1 - t_2| \leq 1 \\ 3, & \text{if } t_2 > t_1 + 1 \end{cases}$$

$$y_n^\star = x_n^\top \gamma_1 - \varepsilon_n, \qquad y_n = \begin{cases} 0 & \text{if} \quad y_n^\star < 0 \\ 1 & \text{if} \quad 0 \leq y_n^\star < \gamma_0 \\ 2 & \text{if} \quad \gamma_0 \leq y_n^\star \end{cases}$$

where we also include an intercept. Then

$$P(y_n = 1 \text{ or } y_n = 2) = \Lambda \left(x_n^\top \gamma_1 \right)$$

$$P(y_n = 2) = \Lambda\left(x_n^\top \gamma_1 - \gamma_0\right)$$

which allows us to construct the following pseudo-likelihood function:

$$\mathcal{Q} = \sum_{y_n=0} \log\left(1 - \Lambda\left(x_{0n}^\top \gamma\right)\right) + \sum_{y_n \neq 0} \log\left(\Lambda\left(x_{0n}^\top \gamma\right)\right) + \sum_{y_n \neq 2} \log\left(1 - \Lambda\left(x_{1n}^\top \gamma\right)\right) + \sum_{y_n=2} \log\left(\Lambda\left(x_{1n}^\top \gamma\right)\right)$$

where

$$x_{0n} = \left(x_n^\top \vdots \mathbf{0}\right)^\top \qquad x_{1n} = \left(x_n^\top \vdots \mathbf{1}\right)^\top \qquad \gamma = \left(\gamma_1^\top \vdots - \gamma_0\right)^\top$$

As before, first and second order derivatives are presented in the Appendix.

In the data and in the simulations, y is defined using the failure time (i.e., the minimum between censoring and retirement dates). Censored observations do not pose problems when the other person in the household is uncensored and retires earlier, since in that case we can determine that retirement happened sequentially. Whereas we can always mark whether retirement was sequential or simultaneous in the simulations, when censoring happens before the retirement of the uncensored partner or both are censored, we cannot determine in the data whether retirement was sequential or not. Since we use the failure time in both data and simulations, censoring introduces the same degree of "noise" in the definition of y in the data and in the simulations.

Overall Auxiliary Model

Our final auxiliary model objective function is then defined by the pseudo–loglikelihood function

$$\log \mathcal{L}_{men}(\alpha_1, \beta_1) + \log \mathcal{L}_{women}(\alpha_2, \beta_2) + \mathcal{Q}(\gamma)$$

and the moment conditions used for estimating the parameters of the structural model are the first order conditions for maximizing this.

As is customary, we choose as our weighting matrix $W = \widehat{J}_0^{-1}$, where

$$\widehat{J}_0 = \widehat{V}\left[\begin{pmatrix} \frac{\partial \log \mathcal{L}_{mn}}{\partial(\alpha_1, \beta_1)} \\ \frac{\partial \log \mathcal{L}_{wn}}{\partial(\alpha_2, \beta_2)} \\ \frac{\partial \mathcal{Q}_n}{\partial \gamma} \end{pmatrix}\right]$$

The (asymptotic) standard errors of the structural estimates are calculated using the formulae in Gourieroux and Monfort (1996).

3 Data

In the United States, full retirement age for those reaching 62 before 2000 was 65 years old. The full retirement age has been increasing ever since until it reaches 67 for those reaching 62 in 2022. Workers who claim benefits early (between ages 62 to 65) have their basic benefit (PIA, primary insurance account) reduced proportionately. Individuals who delay retirement receive increases in benefits for every month of delayed retirement before age 70. (The rate of increase rose gradually until reaching 8 percent for year of delayed retirement in 2005.) Those claiming early retirement are also subject to an earnings test whereby half of the earnings above a certain threshold are withheld. Most of the lost earnings are treated as delayed receipt. (Until 2000, recipients were also subject to an earnings test during the first five years of retirement.) Aside from the OASDI (Old Age, Survivors, and Disability Insurance) program, the SSA also administers the SSI (Supplemental Security Income) program, which provides assistance to individuals age 65 or older as well as disabled. The entitlement level is unrelated to previous work earnings and is based on the individual or couple's income and net worth.

We estimate the model using eight waves of the Health and Retirement Study (every two years from 1992 to 2006) and keep households where at least one individual was 60 years-old or more. Retirement is observed at a monthly frequency. We classify as retired a respondent who is not working, and not looking for work and if there is any mention of retirement through the employment status or the questions that ask the respondent whether he or she considers him or herself retired.[9] To avoid left-censoring, selected households also had both partners working at the initial period. Right-censoring occurs when someone dies or has his or her last interview before the end of the survey. We excluded individuals who were part of the military. Finally, we exclude households with multiple spouses and/or

[9]Specifically we use the classification provided by the variable `RwLBRF`.

couples throughout the period of analysis, couples with conflicting information over marital status or other joint variables and couples of same gender. This leaves us with 1,469 couples. Of those, 384 couples have both husband and wife's uncensored retirement dates. Among the uncensored couples, 33 couples ($\approx 8.6\%$) retire jointly.[10] Figure 5 plots the retirement month of the husbands against the retirement month of the wives for those couples whose retirement month is uncensored for both spouses (January, 1931 is month 1). The points along the 45-degree line are the joint retirements.

We condition covariates on the first "household year": when the oldest partner reaches 60 years-old.[11] The covariates we use are:

1. the age difference in the couple (husband's age - wife's age in years);
2. dummies for race (non-Hispanic black, Hispanic and other race with non-Hispanic whites as the omitted category);
3. dummies for education (high school or GED, some college and college or above with less than high school as the omitted category);
4. indicators of region (NE, SO, and WE with MW or other region as omitted category);
5. self-reported health dummies (good health, very good health, with poor health as the omitted category);
6. an indicator for whether the person has health insurance;
7. the total health expenditure per individual in the previous 12 months for the first two waves and the previous 2 years for the subsequent years[12] (inflation adjusted using the CPI to Jan/2000 dollars);

[10] There are 540 additional couples with only one censored spouse. If those are presumed to have retired sequentially, the proportion of joint retirements among couples with at most one censored spouse is 3.5%. Taking into account the remaining households where both individual retirement dates are censored, would place the proportion of simultaneous retirements somewhere between 2.2% (if all additional households are assumed to retire sequentially) and 39% (if all additional households are assumed to retire simultaneously).

[11] We take the measurements from the first interview after the oldest spouse turns 60.

[12] We use the transformation $\text{sgn}(\text{total health expenditure}) \times \sqrt{\text{total health expenditure}}$. This transforma-

8. indicators for whether the person had a defined contribution (DC) or defined benefit (DB) plan; and

9. financial wealth (inflation adjusted using the CPI to Jan/2000 dollars).[13] This measure includes value of checking, savings accounts, stocks, mutual funds, investment trusts, CD's, Government bonds, Treasury bills and all other savings minus the value of debts such as credit card balances, life insurance policy loans or loans from relatives. It does not include housing wealth or private pension holding.

In Table 2, we present an overview of intra-household differences. Most of the couples marry within their own race but there is substantial variation in term of education. Many couples report different health statuses and in accordance there is substantial difference in health expenditures. There are also differences with respect to insurance and pension ownership. Figure 6 presents the Kaplan-Meier estimates for the retirement behavior in our sample (measured in year of retirement).

4 Results

We now present our estimation results using monthly data on retirement in couples. The discount rate ρ is set to 0.004 per month (i.e., 5% per year) and the threat points are set at the 0.6 times the utility level they would have obtained if his or her partner never retired.[14] The number of simulations in each set of estimates is $R = 10$. Since we cannot detect any visible discontinuity or kink in those plots, we assume that $Z_i(\cdot)$ is smooth. We

tion is in the spirit of a logarithmic transformation of positive variables and implies that large quantities have a decreasing effect. In contrast to a log transformation, it allows us to handle zeroes.

[13]For financial wealth we use the transformation sgn(financial wealth)$\times\sqrt{\text{financial wealth}}$. This transformation is in the spirit of a logarithmic transformation of positive variables and implies that large quantities have a decreasing effect. In contrast to a log transformation, it allows us to handle negative numbers. It is concave for positive values and convex for negative ones. In the computations, we also divide by 10^6 to avoid overflow.

[14]In our estimations, we experimented with other multiples of this scaled between 0 and 1 as well. See the discussion at the end of this section.

assume that $Z_i(t) = t^{\theta 1i}$ implying a Weibull baseline hazard for a model with $\delta = 1$. In our baseline specifications, utility flows while in the labor force are drawn from independent unit exponentials ($K_i \sim \exp(1)$). (In Table 6, we also allow for dependence between K_1 and K_2 using copulas.) Finally, we take $\varphi_i(\mathbf{x}_i) = \exp(\theta_{2i}^{\top}\mathbf{x}_i)$.

Tables 3 and 4 present our estimates. Results are very robust across covariate specifications. There is positive duration dependence: retirement is more likely as the household ages. Age differences tend to increase the retirement hazard for men and decrease it for women. Since men are typically older and we count "family age" from the 60th year of the oldest partner, a larger age difference implies that the wife is younger at time zero and less likely to retire at any "family age" than an older woman (i.e., a similar wife in a household with a lower age difference). Both non-white men and women have lower retirement hazard than non-Hispanic whites, though only Hispanics' coefficients tend to be significant. The hazard of a hispanic woman is about $0.666 (= \exp(-0.407))$ of a white woman's. The hazard of a hispanic man is about $0.678 (= \exp(-0.389))$ of a white man's.

More educated women, but especially those with high school or GED and in some covariate specifications with college, seem to retire earlier than those without high school, but the coefficients on those categories are not statistically significant. For men, college-educated husbands retire later than all other categories and the association is statistically significant. There is some evidence that high school graduates retire earlier but the effect is numerically small and statistically insignificant. There is some evidence that husbands in the Northeast retire earlier whereas those in the South and West retire later than those in the Midwest. The only statistically significant coefficients are those associated with the South though. Geographical region does not seem to play a statistically significant role for women. Furthermore, depending on the covariate specification, Northeast and Southern women have a lower or higher hazard than those in the Midwest. Western wives do seem to retire earlier in all covariate specifications, but then again standard errors are quite imprecise.

Self-reported health lowers the hazard with healthier people retiring later than those in poor health. Only the female coefficient on "good health" is significant in some of the specifications nonetheless. Having health insurance increases the hazard for both husbands and wives, though not in a statistically significant way. Total health expenditures increase the hazard for husband, but lowers it for the wife. (Though health expenditures tend increase as self-reported health decreases, we conjecture this captures health deterioration not completely relected in self-reported variables.) Having a defined benefit contribution pension plan increases the probability of retirement for both genders in a numerically and statistically significant manner. A defined contribution plan affects negatively the male hazard of men but not the female. Wealthier women tend to retire earlier, but financial wealth does not affect the hazard of men significantly.

The interaction parameter ranges from 1.134 to 1.085 across our various specifications. In terms of our model, this means that the utility flow of retirement increases by around 10% when one's partner retires. In terms of the effect on the hazard rate of retirement, this corresponds to about 40% of the effect of having a defined benefit plan for the men.

We have also added spousal variables as covariates to the last specification. Those variables were: dummies for "very good health" and "good health" and dummies for defined benefit and defined contribution pensions. In the simultaneous duration model, the coefficient for the dummy on "West" is now barely significant at 10% for wives, but the coefficient estimates on the remaining variables are essentially the same as in the tables. For males, the spousal coefficients are statistically insignificant at usual levels. For females, only the coefficient on a defined benefit pension plan for the spouse is statistically significant. The absence of an effect of spousal health is in line with previous findings in the literature (e.g., Coile (2004a)). The effect of a husband having a defined benefit plan on a woman's duration (0.324) is comparable with that of the woman herself having a defined benefit pension plan, which is 0.399 once we include spousal covariates. In contrast, the point estimate of the effect of a wife having a defined benefit pension plan on the man's duration (-0.072) is negative, much lower in magnitude and statistically insignificant, when compared to that of the man himself having a defined benefit plan, which is 0.262 once we include spousal covariates.

In Figure 7, we verify the robustness of our estimates to different threat point levels. As mentioned previously, we set $A_i, i = 1, 2$, equal to 0.6 of the utility spouse i would obtain if spouse j never retired. In the graph we plot 95% confidence intervals and point estimates of δ for various proportions of the utility one would get in case the partner were not to retire in the third specification from Tables 3 and 4. As seen from the plots, the confidence set bounds that are at least about 1.03 (for the lower bounds) and at most 1.17 (for the upper bounds). The point estimates hover around 1.11, which is the estimate presented in our main tables.

Because the differential utility from joint retirement may depend on household characteristics, we also split our estimation into households where husband and wife are within 3 years of age apart and households where their age difference is greater than 3 years. The results are presented in Table 5 for the covariates used in specification in Tables 3 and 4. As expected, for households closer in age, δ is higher (1.102 versus 1.083). Interestingly, retirement timing for wives in those households responds more to health conditions and education than in the baseline specification on Table 3. The hazard is also comparatively higher for nonhispanic black wives than in the baseline specification. For households father apart in age, hispanic wives retire later than in the baseline specification and college seems to decrease the retirement hazard while it raises it in the baseline specification. With respect to husbands, those in households closer in age now seem to respond less to a college education, whose coefficient is also less precisely estimated.[15]

We also allow for the possibility of correlation between the unobserved variables K_1 and K_2. We still assume that their marginal distributions are unit exponentials, but assume that any dependence is captured by a Clayton-Cuzick copula function (see Clayton and Cuzick (1985)). More precisely, we suppose that the joint CDF for K_1 and K_2 is given by:

$$F_{K_1,K_2}(k_1, k_2; \tau) = K(1 - \exp(-k_1), 1 - \exp(-k_2); \tau),$$

where

$$K(u, v; \tau) = \begin{cases} \left(u^{-\tau} + v^{-\tau} - 1\right)^{-1/\tau} & \text{for} \quad \tau > 0 \\ uv & \text{for} \quad \tau = 0. \end{cases}$$

Our previous specifications are a special case where $\tau = 0$. When $\tau > 0$, there is positive

[15]The coefficients on `Other race` is quite different across subsamples. The proportion of individuals of other race is nonetheless small in both: around 3.5% among couples more than 3 years of age apart, and only 1.5% among those closer in age.

dependence between variables K_1 and K_2. This copula is typically used to introduce dependence in the duration literature (see, e.g., Heckman and Honoré (1989)).[16] To estimate the model we augment our auxiliary models with the correlation in failure times (including censored observations in both data and simulation moments.) We calculate the correlation between the residuals from a regression of (log) failure time on all covariates for husband and wife. (An alternative is to use the residuals from regressions on spouse-specific variables. The asymptotic distribution for the correlation in this case would nonetheless depend on nuisance parameters (i.e., the regression coefficients), which is not the case if we use the same set of covariates for husband and wife.) The estimates from this specification are presented in Table 6 for the third covariate specification in Tables 3 and 4. Regressor coefficient estimates are slightly higher for education variables but virtually unchanged. The logarithm of τ is estimated at -0.358, which would roughly imply a Kendall's rank correlation coefficient of 0.259. (Kendall's rank correlation for the Clayton-Cuzick copula is equal to $\tau/(2+\tau)$.) Finally, the interaction parameter is estimated at 1.047, compared to 1.115 in our baseline specification (see Tables 3 and 4). (The p-value for the null hypothesis that $\delta \geq 1$ is equal to 0.96.)

Finally, to evaluate whether joint retirement could be an outcome from a common shock, as opposed to the interaction between husband and wife, we compare the average proportional change in financial wealth preceding retirement years for households that retire simultaneously and households that retire sequentially. Financial assets for those who end up retiring simultaneously are much more stable than for couples who retire sequetially: the average relative change in financial wealth across survey waves preceding retirement is 3.197 for those who retire simultaneously versus 8.222 for those households who retire sequentially. Furthermore, there is no discernible statistical difference between the average change in financial assets from survey wave to survey wave for these two groups. Consequently, it is unlikely that shocks to financial wealth explain the joint retirement decision in our sample.

[16] For a fixed value of τ, we can easily draw from this joint CDF by sampling K_1 from a unit exponential and letting K_2 be $F^{-1}(\cdot|K_1;\tau)$ evaluated at a draw from uniform distribution on $[0,1]$.

COUNTERFACTUALS: TO BE ADDED.

5 Concluding Remarks

We have presented a new model that nests the usual generalized accelerated failure time models, but accounts for joint termination of spells and is built upon an economic model of joint decision making. We applied the model to retirement of husband and wife.

Appendix

AUXILIARY MODELS FOR INDIRECT INFERENCE

Log-likelihood Derivatives: Weibull Model

$$\frac{\partial \log \mathcal{L}_i}{\partial \alpha_i} = \sum_{n=1}^{N} (1 - c_{i,n}) \left(\frac{1}{\alpha_i} + \log(t_{i,n}) \right) - \sum_{n=1}^{N} t_{i,n}^{\alpha_i} \log(t_{i,n}) \exp\left(x'_{i,n}\beta_i\right)$$

$$\frac{\partial \log \mathcal{L}_i}{\partial \beta_i} = \sum_{n=1}^{N} (1 - c_{i,n}) x_{i,n} - \sum_{n=1}^{N} t_{i,n}^{\alpha_i} \exp\left(x'_{i,n}\beta_i\right) x_{i,n}$$

$$\frac{\partial^2 \log \mathcal{L}_i}{\partial \alpha_i^2} = -\sum_{n=1}^{N} (1 - c_{i,n}) \frac{1}{\alpha_i^2} - \sum_{n=1}^{N} t_{i,n}^{\alpha_i} \log(t_{i,n})^2 \exp\left(x'_{i,n}\beta_i\right)$$

$$\frac{\partial^2 \log \mathcal{L}_i}{\partial \alpha_i \partial \beta'_i} = -\sum_{n=1}^{N} t_{i,n}^{\alpha_i} \log(t_{i,n}) \exp\left(x'_{i,n}\beta_i\right) x_{i,n}$$

$$\frac{\partial^2 \log \mathcal{L}_i}{\partial \beta_i \partial \beta'_i} = -\sum_{n=1}^{N} t_{i,n}^{\alpha_i} \exp\left(x'_{i,n}\beta_i\right) x_{i,n} x'_{i,n}$$

To impose $\alpha_i > 0$ in our computations we parameterize $\alpha_i = \exp(\theta)$. Then,

$$\frac{\partial \log \mathcal{L}_i}{\partial \theta} = \frac{\partial \log \mathcal{L}_i}{\partial \alpha_i} \frac{\partial \alpha_i}{\partial \theta} = \left(\sum_{n=1}^{N} (1 - c_{i,n}) \left(\frac{1}{\alpha_i} + \log(t_{i,n}) \right) - \sum_{n=1}^{N} t_{i,n}^{\alpha_i} \log(t_{i,n}) \exp\left(x'_{i,n}\beta_i\right) \right) \alpha_i$$

$$\frac{\partial \log \mathcal{L}_i}{\partial \beta_i} = \sum_{n=1}^{N} (1 - c_{i,n}) x_{i,n} - \sum_{n=1}^{N} t_{i,n}^{\alpha_i} \exp\left(x'_{i,n}\beta_i\right) x_{i,n}$$

$$
\begin{aligned}
\frac{\partial^2 \log \mathcal{L}_i}{\partial \theta^2} &= \frac{\partial}{\partial \theta}\left(\frac{\partial \log \mathcal{L}_i}{\partial \alpha_i}\frac{\partial \alpha_i}{\partial \theta}\right) \\
&= \frac{\partial^2 \log \mathcal{L}_i}{\partial \alpha_i^2}\left(\frac{\partial \alpha_i}{\partial \theta}\right)^2 + \frac{\partial \log \mathcal{L}}{\partial \alpha_i}\frac{\partial^2 \alpha_i}{\partial \theta^2} \\
&= \left(-\sum_{n=1}^{N}(1-c_{i,n})\frac{1}{\alpha_i^2} - \sum_{n=1}^{N} t_{i,n}^{\alpha_i}\log\left(t_{i,n}\right)^2 \exp\left(x_{i,n}'\beta_i\right)\right)\alpha_i^2 \\
&\quad - \left(\sum_{n=1}^{N}(1-c_{i,n})\left(\frac{1}{\alpha_i} + \log\left(t_{i,n}\right)\right) - \sum_{n=1}^{N} t_{i,n}^{\alpha_i}\log\left(t_{i,n}\right)\exp\left(x_{i,n}'\beta_i\right)\right)\alpha_i \\
\frac{\partial^2 \log \mathcal{L}_i}{\partial \theta \partial \beta_i'} &= \frac{\partial^2 \log \mathcal{L}_i}{\partial \alpha_i \partial \beta_i'}\frac{\partial \alpha_i}{\partial \theta} = \left(-\sum_{n=1}^{N} t_{i,n}^{\alpha_i}\log\left(t_{i,n}\right)\exp\left(x_{i,n}'\beta_i\right)x_{i,n}\right)\alpha_i \\
\frac{\partial^2 \log \mathcal{L}_i}{\partial \beta_i \partial \beta_i'} &= -\sum_{n=1}^{N} t_{i,n}^{\alpha_i}\exp\left(x_{i,n}'\beta_i\right)x_{i,n}x_{i,n}'
\end{aligned}
$$

Pseudo-likelihood Derivatives: Ordered Model

$$
\frac{\partial \mathcal{Q}}{\partial \gamma} = \sum_n \left[\left(1\left\{y_n \neq 0\right\} - \Lambda\left(x_{0n}^\top \gamma\right)\right)x_{0n} + \left(1\left\{y_n = 2\right\} - \Lambda\left(x_{1n}^\top \gamma\right)\right)x_{1n}\right]
$$

$$
\frac{\partial^2 \mathcal{Q}}{\partial \gamma \partial \gamma^\top} = -\sum_n \left[\left(\left(1-\Lambda\left(x_{0n}^\top \gamma\right)\right)\Lambda\left(x_{0n}^\top \gamma\right)\right)x_{0n}x_{0n}^\top + \left(\left(1-\Lambda\left(x_{1n}^\top \gamma\right)\right)\Lambda\left(x_{1n}^\top \gamma\right)\right)x_{1n}x_{1n}^\top\right]
$$

ADDITIONAL DERIVATIONS FOR IDENTIFICATION DISCUSSION

Here we provide details for the effect of δ on the retirement date second spouse to retire when there is sequential retirement. First, note that when $t_1 \approx 0$, applying the Implicit Function Theorem to the FOC for t_2 (see equation (2)) gives

$$
\frac{dt_2}{d\delta} = -\left[\frac{\frac{\partial^2 I}{\partial t_2 \partial \delta}\times(II) + \frac{\partial II}{\partial \delta}\times\frac{\partial I}{\partial t_2} + \frac{\partial^2 II}{\partial t_2 \partial \delta}\times(I) + \frac{\partial I}{\partial \delta}\times\frac{\partial II}{\partial t_2}}{\frac{\partial^2 I}{\partial^2 t_2}\times(II) + \frac{\partial II}{\partial t_2}\times\frac{\partial I}{\partial t_2} + \frac{\partial^2 II}{\partial^2 t_2}\times(I) + \frac{\partial I}{\partial t_2}\times\frac{\partial II}{\partial t_2}}\right], \tag{4}
$$

where (I) and (II) are defined as in equation (2). The various terms can be signed as shown below:

$$
\begin{aligned}
&\tfrac{\partial I}{\partial \delta} = \varphi_1 \tilde{Z}_1(t_2) > 0 && \tfrac{\partial II}{\partial \delta} = \varphi_2 \tilde{Z}_2(t_2) > 0 \\
&\tfrac{\partial I}{\partial t_2} = Z_1(t_2)e^{-\rho t_2}\varphi_1(1-\delta) < 0 && \tfrac{\partial II}{\partial t_2} = k_2 e^{-\rho t_2} - Z_2(t_2)\varphi_2\delta e^{-\rho t_2} > 0 \\
&\tfrac{\partial^2 I}{\partial t_2 \partial \delta} = -Z_1(t_2)e^{-\rho t_2}\varphi_1 < 0 && \tfrac{\partial^2 II}{\partial t_2 \partial \delta} = k_2 e^{-\rho t_2} - Z_2(t_2)\varphi_2\delta e^{-\rho t_2} > 0 \\
&\tfrac{\partial^2 I}{\partial t_2^2} = Z_1'(t_2)e^{-\rho t_2}\varphi_1(1-\delta) < 0 && \tfrac{\partial^2 II}{\partial t_2^2} = -\rho e^{-\rho t_2}(k_2 - Z_2(t_2)\varphi_2\delta) - Z_2'(t_2)e^{-\rho t_2} < 0.
\end{aligned}
$$

These and the fact that $(I) \geq 0$ and $(II) \geq 0$ imply that the denominator in expression (4) is *strictly* negative. To see that the numerator is also negative notice that

$$\lim_{\delta \to 1} \left[\frac{\partial II}{\partial \delta} \times \frac{\partial I}{\partial t_2} + \frac{\partial I}{\partial \delta} \times \frac{\partial II}{\partial t_2} \right] = \varphi_1 \tilde{Z}_1(t_2)[k_2 - Z_2(t_2)\varphi_2] = 0$$

where the last equality follows because $k_2 = Z_2(t_2)\varphi_2$ at the optimally chosen t_2 when $\delta = 1$. Since

$$\frac{\partial}{\partial \delta} \left[\frac{\partial II}{\partial \delta} \times \frac{\partial I}{\partial t_2} + \frac{\partial I}{\partial \delta} \times \frac{\partial II}{\partial t_2} \right] = -\varphi_1 \varphi_2 \left(Z_1(t_2)\tilde{Z}_2(t_2) + Z_2(t_2)\tilde{Z}_1(t_2) \right) e^{-\rho t_2} < 0,$$

it follows that

$$\frac{\partial II}{\partial \delta} \times \frac{\partial I}{\partial t_2} + \frac{\partial I}{\partial \delta} \times \frac{\partial II}{\partial t_2} < 0.$$

The other two remaining terms in the numerator are negative, which then implies that the numerator is negative. Consequently, (4) is negative: larger values of δ lead to earlier retirement by the second agent (i.e., lower t_2). Because $t_1 \approx 0$, I (and, consequently, t_2) will not depend on k_1. Having identified $Z_i(\cdot)$, $\varphi_i(\cdot)$ and the marginal distribution of K_2, this allows one to identify δ.

References

An, M., B. Christensen, and N. Gupta (2004): “Multivariate Mixed Proportional Hazard Modeling of the Joint Retirement of Married Couples,” *Journal of Applied Econometrics*, 19(6), 687–704.

Banks, J., R. Blundell, and M. Casanova Rivas (2007): “The Dynamics of Retirement Behavior in Couples: Reduced-Form Evidence from England and the U.S.,” Working Paper, University College London (UCL).

Binmore, K., A. Rubinstein, and A. Wolinsky (1986): “The Nash Bargaining Solution in Economic Modelling,” *Rand Journal of Economics*, 17, 176–188.

Blau, D. (1998): “Labor Force Dynamics of Older Married Couples,” *Journal of Labor Economics*, 16(3), 595–629.

Blau, D., and D. Gilleskie (2004): “Health Insurance and Retirement of Married Couples,” Working Paper, University of North Carolina at Chapel Hill.

Browning, M., P.-A. Chiappori, and V. Lechene (2006): “Collective and Unitary Models: A Clarification,” *Review of Economics of the Household*, 4(1), 5–14.

Burtless, G. (1990): “Comment on M. Hurd’s ‘The Joint Retirement Decisions of Husbands and Wives’,” in *Issues in the Economics of Aging*, ed. by D. A. Wise. Chicago: University of Chicago Press.

Casanova Rivas, M. (2010): “Happy Together: A Structural Model of Couples’ Joint Retirement Choices,” Working Paper, University of California at Los Angeles (UCLA).

Chiappori, P.-A. (1992): “Collective Labor Supply and Welfare,” *Journal of Political Economy.*

Chiappori, P.-A., O. Donni, and I. Komunjer (2012): “Learning from a Piece of Pie,” *The Review of Economic Studies*, 79(1), 162–195.

Clayton, D., and J. Cuzick (1985): "Multivariate Generalizations of the Proportional Hazards Model," *Journal of Royal Statistical Society Series B*, 34, 187–220.

Coile, C. (2004a): "Health Shocks and Couples' Labor Supply Decisions," NBER Working Paper 10810.

——— (2004b): "Retirement Incentives and Couples' Retirement Decisions," *Topics in Economic Analysis and Policy*, 4(1).

de Paula, A. (2009): "Inference in a Synchronization Game with Social Interactions," *Journal of Econometrics*, 148(1), 56–71.

Gallant, R., and G. Tauchen (1996): "Which moments to match?," *Econometric Theory*, 12(4).

Gourieroux, C., and A. Monfort (1996): *Simulation-Based Econometric Methods*. Oxford: Oxford University Press.

Gourieroux, C., A. Monfort, and E. Renault (1993): "Indirect Inference," *Journal of Applied Econometrics*, 8, S85–S118.

Gustman, A., and T. Steinmeier (2000): "Retirement in Dual-Career Families: A Structural Model," *Journal of Labor Economics*, 18(3), 503–545.

——— (2004): "Social Security, Pensions and Retirement Behaviour with the Family," *Journal of Applied Econometrics*, 19, 723–737.

——— (2009): "Integrating Retirement Models," NBER Working Paper 15607.

Heckman, J. J., and B. E. Honoré (1989): "The Identifiability of the Competing Risks Model," *Biometrika*, 76, 325–330.

Honoré, B., and A. de Paula (2010): "Interdependent Durations," *Review of Economic Studies*, 77, 11381163.

Hurd, M. (1990): "The Joint Retirement Decisions of Husbands and Wives," in *Issues in the Economics of Aging*, ed. by D. A. Wise. Chicago: University of Chicago Press.

Magnac, T., J.-M. Robin, and M. Visser (1995): "Analysing Incomplete Individual Employment Histories Using Indirect Inference," *Journal of Applied Econometrics*, 10, S153–S169.

Manski, C. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60(3), 531–542.

Marshall, A., and I. Olkin (1967): "A Multivariate Exponential Model," *Journal of the American Statistical Association*, 62, 30–44.

Michaud, P., and F. Vermeulen (2011): "A Collective Labor Supply Model with Complementarities in Leisure: Identification and Estimation by Means of Panel Data," *Labour Economics*, 18, 159–167.

Michaud, P.-C. (2003): "Labor Force Participation Dynamics and Social Security Claiming Decisions of Elderly Couples," Working Paper, CentER Tilburn University and IZA.

Nash, J. (1950): "The Bargaining Problem," *Econometrica*, 18, 155–162.

Ridder, G. (1990): "The Non-Parametric Identification of Generalized Accelerated Failure-Time Models," *Review of Economic Studies*, 57.

Smith, A. (1993): "Estimating nonlinear time series models using vector-autoregressions: Two approaches," *Journal of Applied Econometrics*, 8, S63–S84.

Zeuthen, F. (1930): *Problems of Monopoly and Economic Warfare*. London: Routledge.

Appendix: Figures and Tables

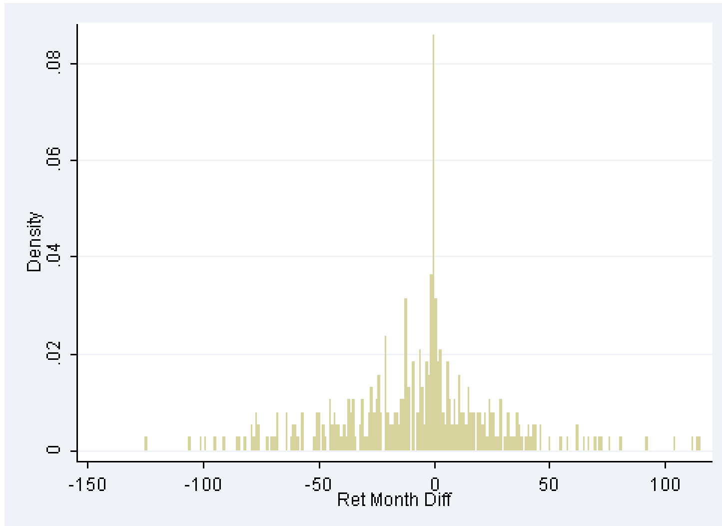

Figure 1: Difference in Retirement Months (Husband-Wife)

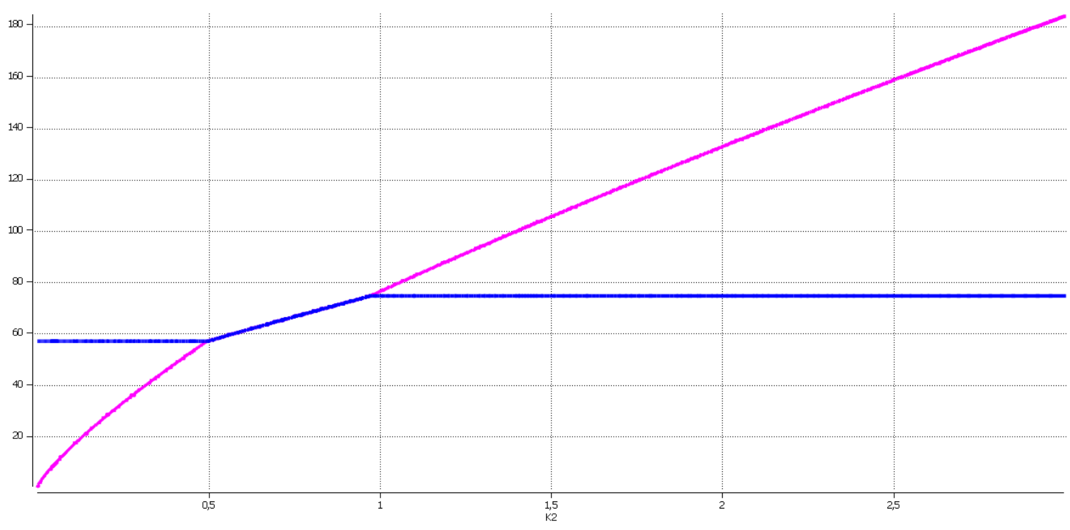

Figure 2: T_1 and T_2 as Functions of K_2 (for K_1 fixed)

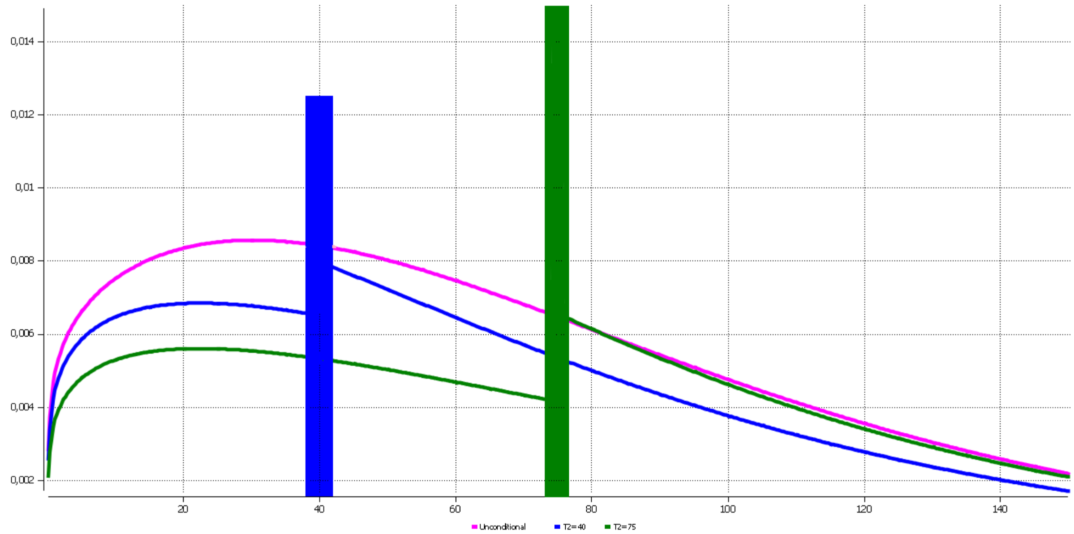

Figure 3: Probability Density Functions for T_1 (given T_2 for conditional pdfs)

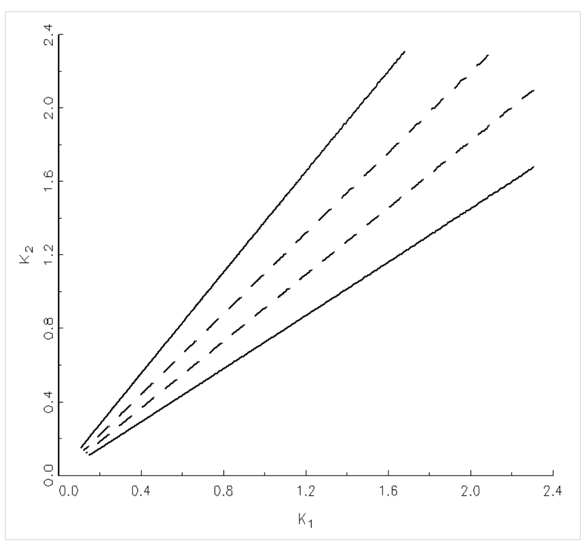

Figure 4: Joint Retirement Regions

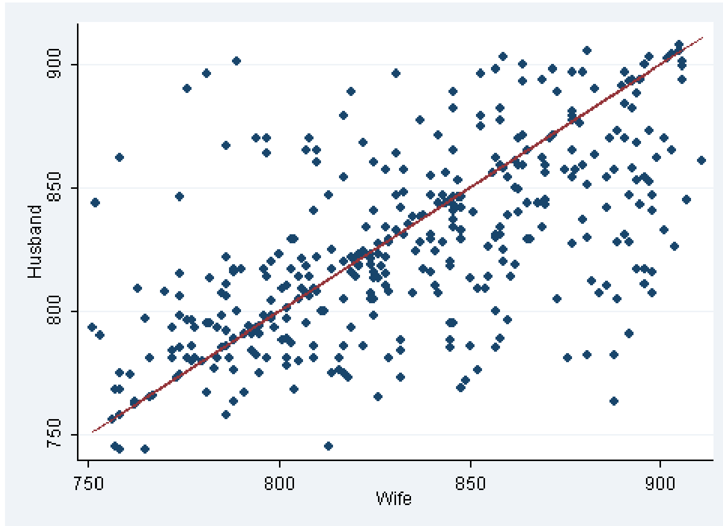

Figure 5: Retirement Months: Husband vs Wife

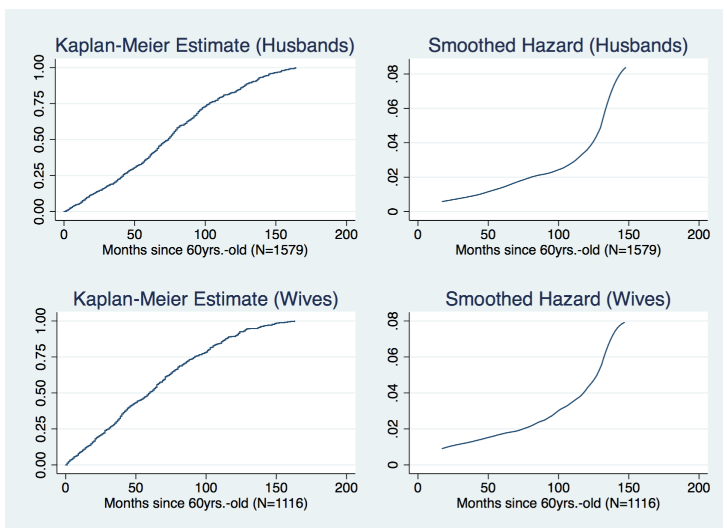

Figure 6: Kaplan-Meier Estimates: Husband and Wife

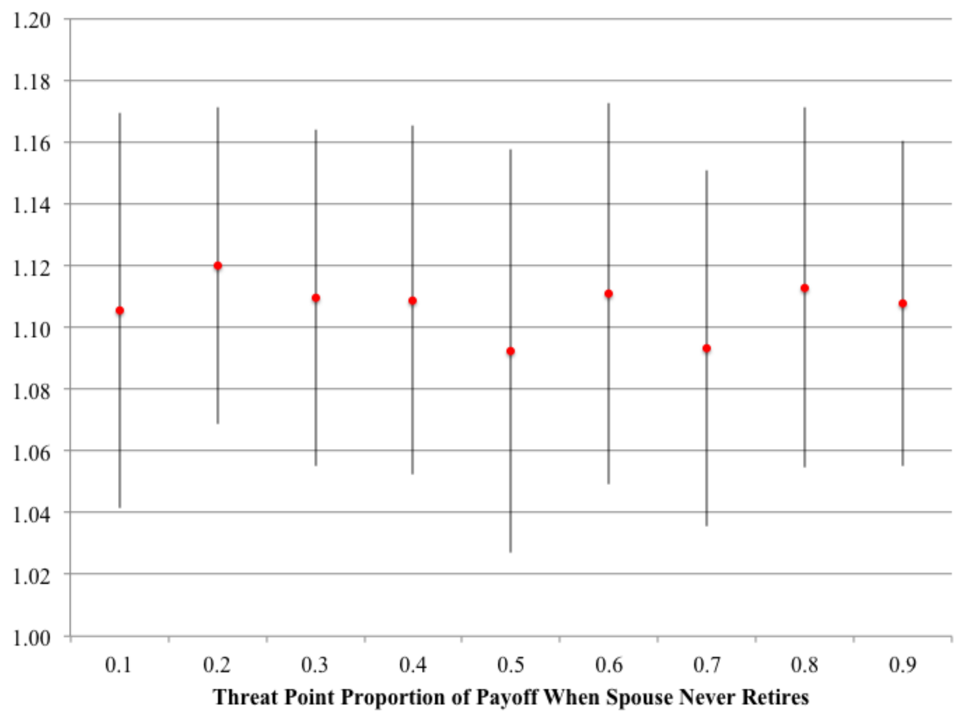

Figure 7: Robustness of $\hat{\delta}$ to Different Threat Point Specifications

Table 1: Summary statistics

	All Observations		Uncensored		Censored	
Variable	**Mean**	**N**	**Mean**	**N**	**Mean**	**N**
Gender	0.50	2938	0.59	1308	0.43	1630
Failure Month	53.60	2938	42.67	1308	62.36	1630
Censored	0.56	2938	0	1308	1	1630
Censoring Month[a]	81.36	2938	105.03	1308	62.36	1630
Age Diff.	3.90	2794	3.62	1268	4.14	1526
Nonhisp. White	0.77	2916	0.809	1298	0.73	1618
Nonhisp. Black	0.10	2916	0.089	1298	0.11	1618
Other Race	0.03	2916	0.023	1298	0.03	1618
Hispanic	0.11	2916	0.079	1298	0.14	1618
$<$ High School	0.19	2916	0.191	1298	0.20	1618
HS or GED	0.36	2916	0.378	1298	0.35	1618
Some College	0.22	2916	0.216	1298	0.22	1618
College or Above	0.22	2916	0.215	1298	0.23	1618
NE	0.17	2916	0.181	1298	0.16	1618
MW	0.24	2916	0.265	1298	0.23	1618
SO	0.42	2916	0.384	1298	0.45	1618
WE	0.17	2916	0.169	1298	0.17	1618
Health Insurance	0.85	2896	0.88	1288	0.83	1608
V Good Health	0.53	2916	0.551	1298	0.52	1618
Good Health	0.31	2916	0.294	1298	0.31	1618
Poor Health	0.16	2916	0.155	1298	0.17	1618
Tot. Health Exp.[b]	8.07×10^3	2436	8.80×10^3	1249	7.30×10^3	1187
Pension (DB)	0.23	2916	0.28	1298	0.18	1618
Pension (DC)	0.22	2916	0.201	1298	0.23	1618
Financial Wealth[b]	94.01×10^3	2938	87.83×10^3	1308	98.97×10^3	1630

a. For those uncensored, the censoring month is the smallest between the last interview or death date. It is used in the simulations for indirect inference.
b. Inflation-adjusted using the CPI to 2000 US Dollars.

Table 2: Intra-Household Differences

	Prop. or Diff.	N of Couples
Same Race (proportion)	0.9523	1447
Same Education (proportion)	0.4731	1469
Same Self-Reported Health (proportion)	0.4755	1447
Health Insurance (both) (proportion)	0.8160	1429
Health Insurance (neither) (proportion)	0.1092	1429
DB Pension (both) (proportion)	0.0636	1447
DB Pension (neither) (proportion)	0.6123	1447
DC Pension (both) (proportion)	0.0560	1447
DC Pension (neither) (proportion)	0.6185	1447
Health Exp. (difference) (US$1,000)	1.900	1208

Only couples with no missing variable. Inflation-adjusted health expenditures in Jan/2000 USD.

Table 3: WIVES' Simultaneous Duration

Variable	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)
δ	1.134 (0.030)	1.118 (0.029)	1.115 (0.027)	1.085 (0.026)	1.088 (0.024)	1.091 (0.025)
θ_1	1.148 (0.039)	1.212 (0.049)	1.210 (0.051)	1.217 (0.054)	1.226 (0.054)	1.227 (0.062)
Constant	-5.627 ** (0.176)	-5.964 ** (0.290)	-5.793 ** (0.279)	-6.011 ** (0.335)	-5.966 ** (0.342)	-6.011 ** (0.347)
Age Diff.	-0.071 ** (0.017)	-0.067 ** (0.019)	-0.068 ** (0.020)	-0.074 ** (0.019)	-0.076 ** (0.019)	-0.077 ** (0.020)
Nonhisp. Black		-0.151 (0.177)	-0.130 (0.160)	-0.049 (0.164)	-0.031 (0.155)	0.027 (0.161)
Other race		-0.355 (0.370)	-0.369 (0.403)	-0.334 (0.296)	-0.304 (0.346)	-0.317 (0.321)
Hispanic		-0.445 * (0.184)	-0.476 ** (0.184)	-0.376 † (0.207)	-0.418 † (0.218)	-0.407 † (0.216)
High school or GED		0.141 (0.168)	0.182 (0.167)	0.211 (0.159)	0.148 (0.156)	0.141 (0.181)
Some college		0.107 (0.171)	0.116 (0.173)	0.076 (0.170)	0.057 (0.175)	-0.009 (0.175)
College or above		0.162 (0.189)	0.176 (0.190)	0.274 (0.185)	0.155 (0.189)	0.067 (0.212)
NE		-0.007 (0.164)	-0.001 (0.158)	-0.083 (0.160)	-0.116 (0.159)	-0.132 (0.161)
SO		-0.002 (0.126)	0.007 (0.125)	0.021 (0.124)	-0.015 (0.119)	-0.019 (0.120)
WE		0.182 (0.151)	0.207 (0.150)	0.167 (0.153)	0.163 (0.153)	0.180 (0.152)
V Good Health			-0.174 (0.139)	-0.207 (0.163)	-0.257 (0.157)	-0.251 (0.164)
Good Health			-0.319 * (0.144)	-0.321 * (0.163)	-0.390 * (0.156)	-0.353 † (0.165)
Health Insurance				0.302 † (0.164)	0.227 (0.179)	0.202 (0.167)
Tot. Health Exp.				-0.155 (0.945)	-0.130 (1.192)	-0.077 (1.027)
Pension (DC)					0.112 (0.124)	0.127 (0.125)
Pension (DB)					0.370 ** (0.117)	0.387 ** (0.116)
Fin. Wealth						0.376 † (0.195)
Obj Func Value	0.024	2.055	2.255	2.349	2.450	2.852
N	1397	1379	1379	1143	1143	1143

Significance levels : † : 10% $*$: 5% $**$: 1%. Significance levels are not displayed for θ_1 nor δ. Ommitted categories are `Non-Hisp. White`, `Less than high school` `Midwest or Other Region`, and `Poor Health` The threat point scale factor is 0.6, $\rho = 0.004$ and $R = 10$.

Table 4: HUSBANDS' Simultaneous Duration

Variable	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)
δ	1.134	1.118	1.115	1.085	1.088	1.091
	(0.030)	(0.029)	(0.027)	(0.026)	(0.024)	(0.025)
θ_1	1.172	1.216	1.218	1.208	1.218	1.221
	(0.039)	(0.044)	(0.045)	(0.039)	(0.045)	(0.042)
Constant	-5.444 **	-5.396 **	-5.341 **	-5.659 **	-5.704 **	-5.699 **
	(0.169)	(0.217)	(0.247)	(0.262)	(0.274)	(0.261)
Age Diff.	0.025 **	0.023 *	0.023 **	0.033 **	0.032 **	0.031 **
	(0.007)	(0.007)	(0.007)	(0.007)	(0.009)	(0.007)
Nonhisp. Black		-0.116	-0.107	-0.144	-0.127	-0.119
		(0.145)	(0.142)	(0.158)	(0.170)	(0.166)
Other race		-0.163	-0.143	-0.059	-0.009	-0.062
		(0.218)	(0.217)	(0.263)	(0.230)	(0.270)
Hispanic		-0.590 **	-0.583 **	-0.437 **	-0.423 **	-0.389 *
		(0.169)	(0.162)	(0.173)	(0.173)	(0.173)
High school or GED		0.047	0.016	0.033	0.050	0.040
		(0.121)	(0.118)	(0.131)	(0.134)	(0.134)
Some college		0.036	0.006	-0.015	-0.018	-0.055
		(0.126)	(0.125)	(0.140)	(0.144)	(0.144)
College or above		-0.300 *	-0.323 *	-0.285 †	-0.257 †	-0.296 *
		(0.129)	(0.127)	(0.139)	(0.141)	(0.144)
NE		0.102	0.111	0.117	0.089	0.084
		(0.114)	(0.113)	(0.126)	(0.120)	(0.124)
SO		-0.230 *	-0.207 *	-0.151	-0.171	-0.170
		(0.100)	(0.099)	(0.112)	(0.110)	(0.110)
WE		-0.110	-0.119	-0.052	-0.078	-0.066
		(0.120)	(0.119)	(0.131)	(0.127)	(0.128)
V Good Health			-0.053	0.028	0.028	0.025
			(0.126)	(0.142)	(0.139)	(0.141)
Good Health			-0.056	0.015	-0.028	-0.039
			(0.127)	(0.146)	(0.142)	(0.145)
Health Insurance				0.226 †	0.198	0.200
				(0.126)	(0.136)	(0.133)
Tot. Health Exp.				1.188 *	1.356 †	1.355 *
				(0.590)	(0.730)	(0.656)
Pension (DC)					-0.204 †	-0.211 *
					(0.106)	(0.107)
Pension (DB)					0.289 **	0.290 **
					(0.104)	(0.104)
Fin. Wealth						0.103
						(0.174)
Obj Func Value	0.024	2.055	2.255	2.349	2.450	2.852
N	1397	1379	1379	1143	1143	1143

Significance levels : † : 10% $*$: 5% $**$: 1%. Significance levels are not displayed for θ_1 nor δ. Ommitted categories are `Non-Hisp. White`, `Less than high school` `Midwest or Other Region`, and `Poor Health` The threat point scale factor is 0.6, $\rho = 0.004$ and $R = 10$.

Table 5: Simultaneous Duration by Age Diff.

	$>$ **3 yrs.**		$\leq$ **3 yrs.**	
	Wife	**Husband**	**Wife**	**Husband**
Variable	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)
δ	1.083 (0.046)		1.1017 (0.0304)	
θ_1	1.200 (0.081)	1.246 (0.049)	1.2657 (0.0551)	1.237 (0.062)
Constant	-6.227 ** (0.460)	-5.429 ** (0.311)	-5.7166 ** (0.3423)	-5.171 ** (0.380)
Nonhisp. Black	-0.569 (0.422)	-0.177 (0.207)	0.3421 † (0.1996)	0.016 (0.242)
Other race	-0.366 (0.442)	0.386 † (0.218)	-1.3068 (0.4810)	-1.129 * (0.549)
Hispanic	-0.982 ** (0.246)	-0.135 (0.209)	-0.0971 (0.2560)	-0.905 ** (0.266)
High school or GED	-0.137 (0.209)	0.041 (0.177)	0.2715 (0.2382)	0.102 (0.187)
Some college	-0.248 (0.206)	-0.135 (0.188)	0.1158 (0.2565)	0.002 (0.199)
College or above	-0.575 * (0.255)	-0.318 † (0.182)	0.5600 * (0.2527)	-0.223 (0.196)
NE	0.250 (0.254)	-0.137 (0.199)	-0.1254 (0.2171)	0.460 (0.148)
SO	0.312 (0.216)	-0.083 (0.163)	-0.1471 (0.1589)	-0.241 (0.157)
WE	0.558 * (0.248)	-0.031 (0.189)	-0.0114 (0.2213)	0.003 (0.167)
V Good Health	-0.158 (0.211)	-0.025 (0.171)	-0.3828 † (0.2097)	-0.261 (0.242)
Good Health	-0.238 (0.231)	0.012 (0.176)	-0.5186 ** (0.2005)	-0.342 (0.251)
Obj Func Value	4.124		2.964	

Significance levels :† : 10% $*$: 5% $**$: 1%. Significance levels are not displayed for θ_1 nor δ. Omitted categories are `Non-Hisp. White`, `Less than high school`, `Midwest or Other Region`, and `Poor Health`. The threat point scale factor is 0.6, $\rho = 0.004$ and $R = 5$.

Table 6: Simultaneous Duration (Copula)

	Wife	Husband
Variable	**Coef.** (Std. Err.)	**Coef.** (Std. Err.)
δ	1.047 (0.026)	
θ_1	1.227 (0.052)	1.226 (0.043)
Constant	-5.792 ** (0.275)	-5.341 ** (0.230)
Age Diff.	-0.067 ** (0.020)	0.023 ** (0.007)
Nonhisp. Black	-0.131 (0.159)	-0.130 (0.160)
Other race	-0.379 (0.341)	-0.191 (0.234)
Hispanic	-0.508 ** (0.176)	-0.587 ** (0.169)
High school or GED	0.212 (0.158)	0.014 (0.120)
Some college	0.157 (0.170)	0.034 (0.131)
College or above	0.220 (0.190)	-0.281 * (0.128)
NE	-0.028 (0.156)	0.087 (0.116)
SO	-0.006 (0.118)	-0.226 * (0.101)
WE	0.205 (0.151)	-0.127 (0.121)
V Good Health	-0.200 (0.145)	-0.064 (0.119)
Good Health	-0.331 * (0.147)	-0.038 (0.127)
$\ln(\tau)$	-0.358 (0.468)	
Obj Func Value	5.588	

Significance levels : † : 10% $\ast$: 5% $\ast\ast$: 1%. Significance levels are not displayed for θ_1, $\ln(\tau)$ nor δ. Omitted categories are `Non-Hisp.White`, `Less than high school`, `Midwest or Other Region`, and `Poor Health`. The threat point scale factor is 0.6, $\rho = 0.004$ and $R = 10$. τ is the parameter for the Clayton-Cuzick copula.