Inferring the Distribution of Households' Duration of Residence From Data on Current Residence Time

Shoshana ANILY and Jacob HORNIK
The Leon Recanati Graduate School of Business, Tel Aviv University, Tel Aviv 69978, Israel (*anily@post.tau.ac.il*) (*hornik@post.tau.ac.il*)

Miron ISRAELI
Soreq Nuclear Research Center, Yavne 70600, Israel (*miron@ndc.soreq.gov.il*)

Estimates of households' expected duration of residence are important to private and public decision makers. The common methods of estimation have been shown to be unreliable. This article presents a measurement framework for estimating total time of residence using five large sets of published government census data on the housing market. By developing a moving-behavior model, the distribution of total residence duration can be estimated from the census data on the age of current residency (i.e., time since moving into current residence). Among other results, we found that the average total residence duration for all U.S. households, 5.5 years, is about half the average age residence time, 10.7 years. This extended intertemporal model provides more reliable estimates for the age and expected duration of occupancy. Therefore, the model better explains and predicts housing-market behavior and also the demand for the many housing-related products and services.

KEY WORDS: Age/interarrival time of a renewal process; Renewal theory; Residential mobility.

The importance of housing in human life, and its special position in public policy decision making, is underscored by the vast amount of literature dealing with the topic. In recent years, however, it has become increasingly clear that many theoretical and practical issues crucial to the study of the housing and related markets cannot be properly understood without valid data on household mobility and duration of residency.

Careful evaluation of housing policy, as well as optimum effectiveness and efficiency of housing programs, requires adequate data on the expected duration of occupancy at the same residence (time between moving into and out of a residence). Reliable data, however, on residential duration are not easy to obtain. Decision makers usually use household interviews to estimate household residence duration (Hempel and Ayal 1977; Soberon-Ferrer and Dardis 1991), thus relying on respondents' willingness to reveal and ability to reconstruct past states. Clearly, these data suffer from the limitations of retrospective questions and subjective estimates common to consumer surveys (Pearson, Ross, and Dawes 1992).

Other sources of data are more objective housing surveys that provide measures of the average age of the current residence (time since moving into current residence). These data are used as a surrogate for computing an estimator on the expected value of duration of occupancy at the same residence (e.g., Harsman and Quigley 1991). Because the distribution of the two quantities is not necessarily the same, neither are their averages. These considerations point to the need for more reliable measures designed to generate up-to-date estimates of households' total residence time.

Expected occupancy duration by household is important information for various private and public decisions and for the understanding of various social and economic phenomena. Notable examples are, first, for predictions of the future housing and housing-related markets. Specifically, the expected time of residence has been shown to have a major influence on the individual housing-buying process (Hempel and Ayal 1977), for predicting occupancy turnover in rental housing units (e.g., Shear 1983), for modeling individuals' residential relocation behavior (Nijkamp, Vanwissen, and Rima 1993), for better understanding landlords' investments and improvements to rental housing [studies (e.g., Read 1991) have shown that landlords tend to economize on maintenance of housing quality when they expect a shorter occupancy period], for investigating the real-estate cycle (Case 1991), and for understanding housing-related financial-service markets such as the mortgage market (Harsman and Quigley 1991).

Second, one of the primary housing decisions made by consumers associated with altering housing consumption and investments is changing residence or undertaking alterations and additions. Therefore, information on consumers' residence duration will assist the understanding and prediction of household expenditure patterns for products that complement and/or substitute for housing. For example, a household may decide, given their expected residence duration, to buy a new washer/dryer combination, a new refrigerator, and a new oven but to delay the previously planned purchase of a new car.

Third, the understanding and prediction of residential mobility and tenure choice and the calculation of mobility

Journal of Business & Economic Statistics
July 1999, Vol. 17, No. 3

indexes (Pickles and Davis 1991) will certainly benefit from improved estimates of occupancy duration. More accurate data will furnish more reliable information on expected local, regional, and national residential movements that have broad implications both for public policy and for commercial decisions such as retail location (Hornik and Feldman 1982).

Fourth, among other factors, expected value of length of residence in the same house is particularly important for explaining general living systems (Reinback and Oliva 1981). These data have been shown to capture the attachments of household members to their neighborhood—their familiarity with the area, their social ties, and their feelings of security—all of which increase with length of residence. Specifically, households with a greater length of residence are less likely to move (Landale and Guest 1985). In many of these studies, although longitudinal data have been used to measure duration of occupancy, the tendency has been to examine relatively small samples due to obvious practical and budget constraints.

Fifth, reliable measures of households' expected value of residence duration can improve estimates of dropout rates from consumer panel and other longitudinal samples—specifically, by reducing attrition bias (due partly to unreliable measures of occupancy duration) in models using consumer panel data (Sharot 1991). In other words, more reliable data will improve procedures to replenish and replace panels to preserve their statistical representation of the research population (Sudman and Ferber 1977; Hornik and Narayana 1982; Golany, Philips, and Rousseau 1991).

In sum, more accurate measures of the distribution function of the residence duration could provide improved estimates of opportunities in the housing and related markets and consequently allow better investment decisions as well as guidelines for local, regional, and national public policy makers (Israeli and Nelson 1992). "Even though the length of time households stay in different dwellings is a key element of the housing decision, it has hardly received any attention at all in the literature" (Ioannides 1987, p. 266). Indeed, Kidd (1977), Harsman and Quigley (1991), and Pickles and Davis (1991) each called for improved measures of occupancy duration. In this article, we develop an estimator for the expected value of total residence duration that appears to be smaller than the commonly used estimator. We present the moving behavior of households as a stochastic model called "renewal process": a renewal occurs whenever a new household moves to a residence. For renewal theory, see Feller (1966) and Ross (1983). Data from the national American Housing Survey for the years 1985, 1987, 1989, 1991, and 1993 were combined with a moving-behavior model to produce summary data on total residence duration. Note that questions related to modeling length of stay arise in various social disciplines and different approaches have been used in the analysis; Gerchak (1984), for example, provided a mathematical model for formalizing such concepts that is based on using measures of residual durations that are commonly used in reliability.

The article is organized as follows: In Section 1 we present the main concepts of renewal theory used thereafter; for homogenous subgroups of the population we provide formulations for calculating the mean, variance, and distribution functions of the total residence duration of a household at the same residence given data on the age of residency in this subgroup. In Section 2 we describe our datasets and the fitted distribution functions to the data within each of the subgroups, and we also provide estimators for the total residence duration for the subgroups and all U.S. households. In Section 3 we discuss our empirical results, and we complete the article with a "Summary and Conclusions" section.

1. RENEWAL THEORY—MAIN CONCEPTS

We choose to approximate the moving behavior of households in a homogenous group as a renewal process, wherein a *renewal* occurs at each epoch that a new household moves into a house. $T_l^k(i)$ for $i \geq 1$ denotes the length of time that the ith household of residence l in group k has lived there. Starting at time 0, let $\{T_l^k(i) \text{ for } 1 \leq i \leq \infty\}$ be a sequence of nonnegative iid random variables, distributed as the random variable T_k, having a common cumulative distribution function F_k, density function f_k, and finite expectation $\mu_k = E(T_k)$. The preceding iid assumption is justified by the fact that we deal here with the occupancy duration of *different households* at the same residence. In renewal theory, the random variable $T_l^k(i)$ is called the *ith interarrival time* of the lth residence in group k. Let $S_l^k(0) = 0$ and $S_l^k(n) = \sum_{i=1}^n T_l^k(i)$ denote the point of time at which the $n+1$st household of residence l in group k moved in, or alternatively the time the nth renewal of residence l in group k has occurred. We use here the common assumption that domiciles are inhabited continuously (e.g., Pickles and Davis 1991; Crone and Mills 1994). For reasons discussed in the Introduction, in obtaining the data, households are not asked directly about the length of time they reside in each of the houses they have lived in. Instead, at a certain point of time t, an interview is conducted in which households are asked only about the year they moved into their current residence. In terms of renewal theory, the information obtained is about the age of the process at time t. Let $N_l^k(t)$ denote the number of renewals that have occurred by time t at residence l in group k (i.e., number of households that have moved into residence l in group k since time 0). The *age* (*lifetime*) of the process at time t of residence l in group k is denoted as $A_l^k(t)$, where $A_l^k(t) = t - S_l^k(N_l^k(t))$. $A_l^k(t)$ represents the length of time since the last renewal that has been observed by time t at the lth residence in group k. Similarly, the *residual waiting time* (*excess life*) of the lth residence in group $k, Y_l^k(t)$ is defined as the length of time from t till the occurrence of the first renewal after t at the lth residence in group k. The pairs $(A_l^k(t), Y_l^k(t))$ for houses $l = 1, 2, \ldots$ in group k are assumed to be independent and identically distributed as $(A_k(t), Y_k(t))$. Note that the independence assumption here is natural because we deal with the age and the residual time with respect to *different houses* and therefore different households. As is well known in *renewal theory,* the distribution of $A_k(t) + Y_k(t)$ is not necessarily identical to the distribution of the interarrival time T_k. Indeed, $A_k(t) + Y_k(t)$ tends to be longer than T_k because t has a greater probability of falling within a large interarrival interval. Most results in renewal theory (e.g., see Ross 1983), deal with the equilibrium distribution of the age and the residual waiting time of the process given the distribution of the interarrival times $T_l^k(i)$. In this study, however, we have data about the age of the process, and by using statistical fitting techniques we estimate the equilibrium distribution of the age of the process. Let G_k and g_k denote the equilibrium cumulative distribution and density functions, respectively, of the age of the process in group k. Moreover, let A_k be a random variable that is distributed according to the density function g_k.

The expected value of the total residence duration in group k, which we denote by $E(T_k)$, is given by

$$\mu_k = E(T_k) = \int_0^\infty (1 - F_k(x))\,dx. \tag{1}$$

We note that, in our context, $E(T_k)$ represents the average occupancy time at a residence in group k. The expected value of the interarrival time $E(T_k)$, for homogenous groups, can be obtained according to Feller (1966) (by differentiating equation XI.4.10) as follows:

$$g_k(x) = \Pr(T_k \geq x)/\mu_k = (1 - F_k(x))/\mu_k; \tag{2}$$

thus,

$$\mu_k = 1/g_k(0). \tag{3}$$

In view of (3), the expected value of the interarrival time, and in our context the expected occupancy duration at a residence in group k, equals the inverse of the equilibrium density function of the age of the process at time 0. As a consequence of Equation (2), because $\Pr(T_k \geq x)$ is a nonincreasing function in x, the equilibrium density function $g_k(x)$ should be also. In addition, given the density function of the age of the process in group k at equilibrium, the cumulative distribution function of the total residence duration in that group may be obtained by

$$1 - F_k(x) = \mu_k g_k(x) \tag{4}$$

and

$$G_k(x) = \frac{\int_0^x (1 - F_k(s))\,ds}{\mu_k}. \tag{5}$$

The second moment of T_k may be computed by using the following equality (see Ross 1983, p. 27, problem 1.1):

$$E(T_k^2) = 2\int_0^\infty x(1 - F_k(x))\,dx. \tag{6}$$

Combining the preceding equation with Equation (4) yields an expression for the variance of the total residence duration in group k, which we denote by $\sigma^2(T_k)$:

$$\begin{aligned}\sigma^2(T_k) &= E(T_k^2) - \mu_k^2 = 2\mu_k \int_0^\infty x g_k(x)\,dx - \mu_k^2 \\ &= \mu_k[2E(A_k) - \mu_k]. \end{aligned} \tag{7}$$

In view of this last equation, because $\sigma^2(T_k)$ must be nonnegative, the mean of the age of the process at equilibrium must be at least as large as half the mean of the interarrival time. In the special case that renewals occur according to a Poisson process, the age of the process at equilibrium is exponentially distributed with the same mean as the interarrival times—refer to the *bus-stop paradox.* In the context of the moving-behavior model, the outcomes are even more surprising: As will be shown, the mean age of the residence time at equilibrium is approximately twice as large as the total residence time.

2. MODEL FORMULATION

2.1 Data

Reports published by the Bureau of Census (U.S. Department of Commerce and the U.S. Department of Housing and Urban Development 1988, 1989, 1991, 1993) include data for housing surveys performed in 1985, 1987, 1989, 1991, and 1993 in which all housing units were revisited. The number of households participating in these surveys ranges from 88 to 95 million. These surveys report, among other data, the 5–10-year period when participating households moved into their current residence, whether their residence is rented or owned, and also its type and its regional location. By using these five datasets, we estimate the equilibrium distribution of the length of time since households moved into their current residence (i.e., the age of the process).

Table 1 presents data obtained from the American Housing Surveys performed in 1985, 1987, 1989, 1991, and 1993. The values shown are the number of households (in thousands) that moved into their current home during the specified time period before the survey. In these and the following tables, ALLHSES means *all households*; we use the index $k = 0$ to represent all households; for example, μ_0 represents the expected intermove time of all households. The surveys contain data about three different partitions of all households into mutually exclusive sets—(1) a partition into *renter* ($k = 1$) versus *owner* ($k = 2$), (2) a partition into *urban* ($k = 3$) versus *rural* ($k = 4$), and (3) a partition according to geographic regions—*Northeast* ($k = 5$), *Midwest* ($k = 6$), *South* ($k = 7$), and *West* ($k = 8$). In addition, we are given data about *Farms* ($k = 9$), which is a subset of the rural-owner category. Unfortunately, the Bureau of Census does not keep records about cross-section groups; as a result, the preceding partitions are not sufficiently fine, and the subgroups obtained are large and consist of a variety of households. We assume in the following analysis, however, that each of these subgroups is homogenous.

We used these datasets to calculate for each group the fraction of households that the surveys found living in their current residences for x years or more. These are estimators for the values $1 - G_k(x)$. We let $\hat{G}_k(x)$ denote the estimator function for $G_k(x)$ $0 \leq k \leq 9$. Values of $1 - \hat{G}_k(x)$ are shown in Table 2.

2.2 Estimation of Current Residence Time and Total Residence Duration for Each Group

According to the notation used in Section 1, $G_k(x)$ rep-

Table 1. Number of Households (in thousands) That Moved Into Their Current Residences During the Time Period Indicated (from American Housing Surveys in years 1985–1993)

	Year moved in									
	1990–1993	*1985–1989*	*1980–1984*	*1975–1979*	*1970–1974*	*1960–1969*	*1950–1959*	*1940–1949*	*Before 1940*	*Total*
					1985 Survey					
ALLHSES		13,762	28,539	15,729	9,221	10,800	6,257	2,563	1,554	88,425
RENTERS		10,181	14,407	3,960	1,645	1,292	463	184	147	32,279
OWNERS		3,581	14,132	11,769	7,576	9,508	5,794	2,379	1,407	56,146
URBAN		11,213	21,797	11,339	6,472	8,155	4,649	1,691	915	66,231
RURAL		2,549	6,743	4,390	2,749	2,646	1,608	872	639	22,196
NE-RGN		1,993	5,679	3,373	2,058	2,608	1,734	779	506	18,730
MW-RGN		3,196	6,986	4,017	2,283	2,859	1,654	684	464	22,143
S-RGN		5,216	9,717	5,147	3,206	3,568	1,957	793	460	30,064
W-RGN		3,357	6,157	3,193	1,674	1,766	912	307	125	17,491
FARMS		82	370	364	245	267	224	169	133	1,854
					1987 Survey					
ALLHSES		33,920	17,436	12,613	7,907	9,711	5,720	2,292	1,289	90,888
RENTERS		20,748	6,462	2,674	1,212	997	369	162	101	32,725
OWNERS		13,172	10,973	9,939	6,695	8,714	5,352	2,130	1,188	58,163
URBAN		26,925	12,750	8,920	5,446	7,223	4,208	1,511	763	67,746
RURAL		6,995	4,686	3,693	2,461	2,488	1,512	781	525	23,141
NE-RGN		5,715	3,795	2,677	1,766	2,369	1,522	679	429	18,952
MW-RGN		7,717	4,300	3,226	1,975	2,510	1,562	628	349	22,267
S-RGN		12,216	5,849	4,191	2,795	3,307	1,791	721	401	31,271
W-RGN		8,273	3,492	2,519	1,371	1,525	845	263	110	18,398
FARMS		254	263	273	230	245	228	132	115	1,740
					1989 Survey					
ALLHSES		44,914	13,535	10,946	7,122	8,883	5,188	2,014	1,081	93,683
RENTERS		25,283	4,146	1,975	951	847	327	155	83	33,767
OWNERS		19,631	9,389	8,972	6,171	8,036	4,861	1,860	997	59,917
URBAN		34,927	9,572	7,591	4,817	6,584	3,826	1,347	628	69,292
RURAL		9,986	3,963	3,355	2,305	2,299	1,362	667	452	24,389
NE-RGN		7,737	3,107	2,292	1,606	2,209	1,447	643	350	19,391
MW-RGN		10,385	3,326	2,801	1,807	2,338	1,364	541	307	22,869
S-RGN		16,171	4,498	3,687	2,537	2,979	1,623	603	330	32,428
W-RGN		10,620	2,605	2,166	1,173	1,357	754	228	94	18,997
FARMS		312	222	276	195	259	195	92	109	1,660
					1991 Survey					
ALLHSES	24,534	27,054	10,613	9,369	6,233	7,933	4,754	1,772	885	93,147
RENTERS	16,908	10,408	2,784	1,363	772	666	260	120	69	33,350
OWNERS	7,626	16,646	7,829	8,006	5,462	7,267	4,494	1,651	816	59,797
URBAN	19,666	19,797	7,351	6,410	4,261	5,861	3,487	1,150	512	68,495
RURAL	4,868	7,257	3,261	2,959	1,973	2,071	1,267	621	373	24,650
NE-RGN	3,897	5,084	2,493	2,041	1,370	1,941	1,301	543	293	18,963
MW-RGN	5,432	6,583	2,591	2,376	1,508	2,048	1,231	494	231	22,594
S-RGN	9,006	9,337	3,584	3,093	2,198	2,689	1,480	523	282	32,192
W-RGN	6,199	6,050	1,945	1,860	1,059	1,255	743	212	79	19,402
FARMS	147	295	205	253	175	207	163	95	88	1,628
					1993 Survey					
ALLHSES	38,106	19,897	8,933	8,385	5,739	7,244	4,173	1,510	737	94,724
RENTERS	23,079	5,767	2,013	1,059	595	583	209	104	63	33,472
OWNERS	15,026	14,130	6,920	7,326	5,144	6,661	3,964	1,406	674	61,251
URBAN	29,649	14,145	6,048	5,611	3,844	5,318	3,079	988	409	69,091
RURAL	8,457	5,752	2,885	2,774	1,895	1,926	1,094	521	328	25,632
NE-RGN	6,268	3,826	2,113	1,748	1,274	1,804	1,145	469	257	18,904
MW-RGN	8,852	4,896	2,167	2,131	1,439	1,881	1,072	396	198	23,032
S-RGN	13,793	6,811	3,006	2,846	2,031	2,434	1,320	482	213	32,936
W-RGN	9,193	4,361	1,647	1,661	996	1,125	637	163	68	19,851
FARMS	195	236	174	208	151	190	130	71	68	1,423

resents the fraction of households in group k found in a census survey to have moved into their current residences less than x years before the survey ($G_k(0) = 0$). We assume here that $G_k(x)$ is ergodic, meaning that it is independent of the date of the survey. This assumption is because data are available only from 1985, which is too short a period to learn about time dependence of the current residence-times distribution. When data over a sufficient number of years are accumulated, we will be able to allow $G_k(x)$ to be nonergodic. Trend-based methods may then be used to

Table 2. Fraction of Households That Moved Into Their Current Residence x Years or More Before the Survey ($1 - \hat{G}_k(x)$)

Year of survey	x	ALLHSE	RENTER	OWNER	URBAN	RURAL	NE-RGN	MW-RGN	S-RGN	W-RGN	FARMS
	0	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
1985	1	.844370	.684590	.936220	.830700	.885160	.893590	.855670	.826500	.808070	.955770
1991	2	.736610	.493014	.872469	.712884	.802515	.794495	.759582	.720241	.680497	.909705
1987	3	.626790	.365990	.773530	.602560	.697720	.698450	.653430	.609350	.550330	.854020
1993	4	.597715	.310498	.754682	.570870	.670061	.668430	.615665	.581218	.536900	.862966
1989	5	.520575	.251251	.672363	.495945	.590553	.601000	.545892	.501326	.440964	.812048
1985	6	.521620	.238270	.684520	.501590	.581370	.590390	.540170	.503290	.456060	.756200
1991	7	.446166	.180930	.594093	.423856	.508114	.526394	.468222	.430200	.368673	.728501
1987	8	.434950	.168530	.584870	.414360	.495220	.498210	.460320	.422310	.360530	.702870
1993	9	.387663	.138205	.523991	.366140	.445654	.466039	.403091	.374423	.317213	.697119
1989	10	.376098	.128469	.515663	.357805	.428062	.440771	.400455	.362619	.303837	.678313
1985	11	.343740	.115590	.474900	.330390	.383580	.410300	.358760	.332090	.273510	.559870
1991	12	.332228	.097451	.463167	.316534	.375822	.394927	.353545	.318868	.268426	.602580
1987	13	.296180	.086810	.413990	.282690	.335640	.356950	.315440	.288290	.223610	.545980
1993	14	.293358	.078065	.411014	.278604	.333099	.354264	.309005	.283155	.234245	.574842
1989	15	.259257	.069980	.365923	.248254	.290500	.322572	.277975	.248921	.189819	.512048
1985	16	.239460	.064620	.339970	.232670	.259730	.300430	.255660	.225450	.177810	.427720
1991	17	.231645	.056582	.329281	.222951	.255781	.287296	.248385	.222788	.172560	.447174
1987	18	.209180	.049780	.298880	.202300	.229290	.263770	.226750	.198910	.149090	.413790
1993	19	.204837	.046427	.291407	.197392	.224875	.261796	.216481	.196745	.150572	.428672
1989	20	.183235	.041816	.262930	.178736	.195990	.239750	.198959	.170686	.128073	.394578
1991	22	.164729	.033433	.237938	.160742	.175740	.215050	.177215	.154510	.117978	.339681
1993	24	.144251	.028651	.207425	.141755	.150944	.194403	.154003	.135080	.100398	.322558
1985	26	.117320	.024600	.170630	.109540	.140520	.161190	.126540	.106770	.076840	.283710
1987	28	.102330	.019310	.149060	.095680	.121780	.138770	.114030	.093150	.066200	.272990
1989	30	.088415	.016732	.128812	.083718	.101726	.125832	.096725	.078821	.056641	.238554
1991	32	.079562	.013463	.116411	.075173	.091724	.112693	.086572	.070980	.053293	.212531
1993	34	.067776	.011233	.098676	.064784	.075804	.098974	.072334	.061179	.043726	.189037
1985	36	.046560	.010250	.067430	.039350	.068080	.068610	.051840	.041680	.024700	.162890
1987	38	.039400	.008040	.057050	.033570	.056440	.058460	.043880	.035880	.020270	.141950
1989	40	.033037	.007048	.047683	.028503	.045881	.051209	.037081	.028771	.016950	.121084
1991	42	.028525	.005667	.041256	.024265	.040325	.044086	.032088	.025006	.014998	.112408
1993	44	.023722	.004989	.033959	.020220	.033123	.038405	.025790	.021102	.011637	.097681
1985	46	.017570	.004550	.025060	.013820	.028790	.027020	.020950	.015300	.007150	.071740
1987	48	.014180	.003090	.020430	.011260	.022690	.022640	.015670	.012820	.005980	.066090
1989	50	.011539	.002458	.016640	.009063	.018533	.018050	.013424	.010176	.004948	.065663
1991	52	.009501	.002069	.013646	.007475	.015132	.015451	.010224	.008760	.004072	.054054
1993	54	.007780	.001882	.011004	.005920	.012797	.013595	.008597	.006467	.003426	.047786

estimate the current residence times distribution in group k at any point in time in the future. Figure 1 presents the data of $1 - \hat{G}_k(x)$ for all groups of households. Data values are spread over two to three orders of magnitude and are presented, therefore, using semilogarithmic scale. It is important to note that the data for all five sets of surveys fit almost perfectly the same curve, which validates the assumption of ergodicity during the years 1985–1993.

We assume that each group is homogeneous with common G_k and F_k functions. The estimator of the equilibrium density function $\hat{g}_k(x)$ for $k \geq 0$, the length of time a household has lived in his current residence (the age of the process), is expressed by

$$\hat{g}_k(x) = -\frac{d}{dx}\,(1 - \hat{G}_k(x)). \tag{8}$$

As can be seen from Figure 1, the values of $\log[1-\hat{G}_k(x)]$ decrease faster for $x < 10$ and for $x > 35$. The slope of this function is the moving-rate factor to the fraction of households that stayed in their current residence at least x years before the survey. It means that the moving rate is higher in the first years after moving into a house and in the group of the households that stayed over 30 years at the same residence.

To estimate the value of $\hat{g}_k(0)$ (and consequently $\hat{\mu}_k$), the values of a_k and b_k were calculated by fitting the data to the function $1 - \hat{G}_k(x) = \exp(-a_k x(1 - b_k x))$ ($a_k > 0$ and $b_k \geq 0$) for $0 \leq x \leq 5$. These functions fit well the moving-rate data in the first years and therefore enable the calculation of $\hat{g}_k(x)$ for $x = 0$. Because $\hat{g}_k(x) = a_k(1 - 2b_k x)\exp(-a_k x(1 - b_k x))$, then $\hat{g}_k(0) = a_k$ and $\hat{\mu}_k = 1/a_k$. In the same region, $(d/dx)\hat{g}_k(x) = -a_k[2b_k + a_k(1 - 2b_k x)^2]\exp(-a_k x(1 - b_k x))$, and therefore $\hat{g}_k(x)$ is nonincreasing. In the region of $5 \leq x \leq 35$, the data can be fitted to functions of the form $1 - \hat{G}_k(x) = d_k \exp(-c_k x)$ ($c_k > 0$ and $d_k > 0$). Similar functions (with different values of the parameters) can be fitted to the data in the region of $x \geq 35$. $\hat{g}_k(x)$ is, therefore, also nonincreasing in these regions.

Let $\hat{A}_k$ denote the estimator for the mean equilibrium age of the process in group k, for $k = 0, \ldots, 9$. In view of the fact that the available data to date enables us to calculate $\hat{G}_k(n)$ for $n = 0, 1, 2, \ldots, 20$ and $n = 22, 24, \ldots, 54$ only, we propose to use a linear interpolation of $\log[1-\hat{G}_k(x)]$ for $x \geq 0$ and to calculate $\hat{A}_k$ numerically using the following

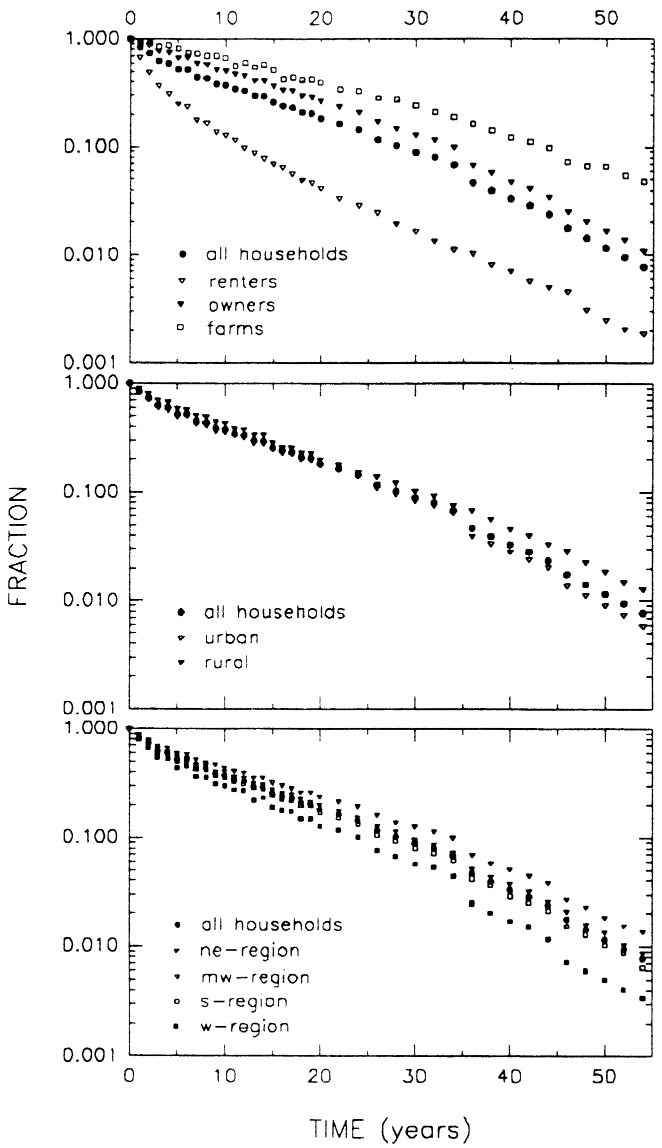

Figure 1. Fraction of Households That Moved Into Their Current Residences x Years or More Before the Survey (observed data).

equation:

$$\hat{A}_k = \int_0^\infty [1 - \hat{G}_k(s)]\, ds. \qquad (9)$$

Let $s(T_k)$ be the estimator for the standard deviation of the average total residence time $\sigma(T_k)$. Based on (7), $s(T_k)$ can be calculated, using the following equation:

$$s(T_k) = \sqrt{\hat{\mu}_k(2\hat{A}_k - \hat{\mu}_k)}.$$

The empirical distribution of $1 - \hat{G}_k(x)$ is not surprising. The process of residential changes has been studied extensively. The prime example remains the seminal work *Why Families Move* (Rossi 1980). Rossi's most often cited contribution was the finding that the function of mobility is "the process by which families adjust their housing to the housing needs that are generated by shifts in family composition that accompany life-cycle change" (Rossi 1980, p. 9). Indeed, many researchers demonstrated that, over their lifetime, households are expected to move frequently in their early years and also change their housing in their late-life-cycle stage. Ioannides (1987), for example, used a panel of 1,257 families to show that the households' propensity to change their place of residency at the early stages is strongly affected by their desire to adjust their housing consumption behavior. In other words, frequent movements in the early household life-cycle stage are caused by movements between rental units (Morrow-Jones 1988), with mobility among renters being almost four times higher than among owners (Ioannides 1987); changes from rental to owned units; and attempts to pursue employment opportunities and household formation and dissolution (Long 1988; Crone and Mills 1994). Residential changes in the late stages of households' life cycle are obviously explained by desertion or dissolution caused by the death of a spouse and by children leaving their parents home (Boehm 1993). Thus, the relationships between households' life history and housing changes have been conceptually related to the framework of survival models.

2.3 Estimated Mean Total/Current Residence Time for All Households

Our objective in this article is to estimate the current/total residence time for a *randomly selected household* in the U.S. population. We note that the current-residence sampling distribution for a randomly selected household is equivalent to that of a randomly selected domicile. The reason for that is that at time t, when the survey is conducted, there is a one-to-one matching between households and domiciles. Therefore, the aggregate current-residence time distribution for a randomly selected household whose membership is unknown is simply the weighted average of the corresponding estimated distributions of the subgroups. To do that, we let n_k denote the number of households in group k. According to our data (see Subsection 2.1), $n_0 = n_1 + n_2, n_0 = n_3 + n_4$, and $n_0 = n_5 + n_6 + n_7 + n_8$. Moreover, let $p_k = n_k/n_0$ for $k = 1, \ldots, 8$ and note that $\sum_{k=1}^{8} p_k = 3$. [Recall, that FARMS ($k = 9$) are a subdivision of the rural-owner category; therefore, we cannot include $\hat{G}_9(x)$ in the weighted average because we do not have data on their complement group—i.e., all nonfarmers.] Let

$$\hat{G}_0(x) = \frac{1}{3}\sum_{k=1}^{8} p_k \hat{G}_k(x) \qquad (10)$$

and

$$\hat{A}_0 = \frac{1}{3}\sum_{k=1}^{8} p_k \hat{A}_k. \qquad (11)$$

Observe that we cannot use Equation (2) to infer from $\hat{G}_0(x)$ about $\hat{F}_0(x)$ because the group of all households is certainly not a homogenous group, and therefore the moving behavior of all households cannot be described as a renewal process. Moreover, as will be explained, the parameter that is the mean total time for a randomly selected domicile is different from the parameter that is the mean total time for a randomly selected household. For example, suppose that the domiciles are divided into two equal-size groups, with constant total residence times of 49 years in the first and 1 year in the second. Then, the average total time for a randomly selected domicile is the average of these two numbers—that is, 25 years. When calculating the mean total time *per household,* however, we must take into account that in the second group the number of individual sojourns is 49 times larger than in the first group, meaning that the proportion of sojourns in the first (second) group is 2% (98%). Thus, the mean total residence time per household is 1.96 years.

Assume that T is a sufficiently large time interval and N is the population size—that is, total number of domiciles in the states (assume for simplicity that this number is fixed). According to our notation, proportion p_1 of domiciles are rented and p_2 of domiciles are owned ($p_1 + p_2 = 1$). The average number of moves out of the rented (owned) domiciles during this time interval is $p_1 NT/\hat{\mu}_1$ ($p_2 NT/\hat{\mu}_2$), and therefore in total the average number of moves within this time interval is $NT(p_1/\hat{\mu}_1 + p_2/\hat{\mu}_2)$; therefore, the average total residence duration of a randomly selected household is the weighted harmonic average of $\hat{\mu}_1$ and $\hat{\mu}_2$:

$$\hat{\mu}_0^1 = \frac{NT}{NT(p_1/\hat{\mu}_1 + p_2/\hat{\mu}_2)} = \frac{1}{p_1/\hat{\mu}_1 + p_2/\hat{\mu}_2}. \quad (12a)$$

We note that $\hat{\mu}_0^1$ is a weighted average of $\hat{\mu}_1$ and $\hat{\mu}_2$; that is, $\hat{\mu}_0^1 = q_1\hat{\mu}_1 + q_2\hat{\mu}_2$, where the weights q_k are the probabilities of getting a type-k sojourn (i.e., a move out of a domicile of type k), which in general are not equal to the probabilities of type-k domicile. To find q_k we have to count the number of sojourns and then see how many of them are of type k. From the preceding, we see that

$$q_k = \frac{p_k/\hat{\mu}_k}{p_1/\hat{\mu}_1 + p_2/\hat{\mu}_2}.$$

Similarly we obtain the following other two estimators for $\hat{\mu}_0$:

$$\hat{\mu}_0^{II} = \frac{1}{p_3/\hat{\mu}_3 + p_4/\hat{\mu}_4} \quad (12b)$$

and

$$\hat{\mu}_0^{III} = \frac{1}{p_5/\hat{\mu}_5 + p_6/\hat{\mu}_6 + p_7/\hat{\mu}_7 + p_8/\hat{\mu}_8}. \quad (12c)$$

Given the information about the three partitions, we estimate the total average number of moves in a population of size N within a time interval of length T as

$$\frac{NT\sum_{k=1}^{8} p_k/\hat{\mu}_k}{3}.$$

Thus, we obtain the following estimator $\hat{\mu}_0$, which is based on the three different partitions:

$$\hat{\mu}_0 = \frac{3}{\sum_{k=1}^{8} p_k/\hat{\mu}_k}. \quad (13)$$

3. CALCULATIONS

3.1 Fitting Procedure and Empirical Results for the Subgroups

The data in Table 2 for $k = 1, \ldots, 9$ and $0 \le x \le 5$ were fitted to the function $1 - \hat{G}_k(x) = \exp(-a_k x(1 - b_k x))$ by a nonlinear least squares regression procedure. The values of a_k and b_k and their asymptotic standard errors calculated for each group are provided in Table 3.

We calculate $\hat{A}_k$ from Equation (9) by interpolating the values of $\log[1 - \hat{G}_k(x)]$ between every two data points as explained in Subsection 2.2. Recall that $\hat{A}_k$ denotes the sample average equilibrium age of the process for subgroup k and $\hat{\mu}_k$ denotes the calculated estimator from the data for the mean total residence duration at the same residence for subgroup $k, 1 \le k \le 9$, as given in (3). Table 3 shows the values of $\hat{\mu}_k$ and of $\hat{A}_k$ for each category $k = 1, \ldots, 9$. $S(T_k)$, the estimators for the standard deviation of the average total residence times $\sigma(T_k)$, and the fractions p_k of all residences in each group are also shown.

Table 3 demonstrates a few more interesting findings: The partition of ALLHSES into RENTERS and OWNERS is the most extreme of all three partitions in terms of the difference between the subgroups in the partition: The estimator $\hat{\mu}_k$ for OWNERS is 533% larger than for RENTERS; $S(T_k)$ for OWNERS is 347% larger than for RENTERS, and the estimator of the coefficient of variation $S(T_k)/\hat{\mu}_k$ is about 50% higher for RENTERS (1.66) than for OWNERS (1.08). We therefore believe that our assumption regarding the homogeneity of these two subgroups is reasonable. As can be expected, the mobility rate of the two categories OWNERS and FARMS is relatively low; their estimator for the coefficient of variation is also the lowest among all subgroups—1.08 and 1.064, respectively.

Table 3. Values of a_k, b_k, and $\hat{\mu}_k$ and their Standard Errors as Well as Values of p_k, $\hat{A}_k$, and $s(T_k)$

Category	p_k	a_k	$s(a_k)$	b_k	$s(b_k)$	$\hat{\mu}_k$	$s(\hat{\mu}_k)$	$\hat{A}_k$	$s(T_k)$
RENTERS	.3594	.408372	.008210	.066039	.004196	2.44875	.04923	4.59622	4.06369
OWNERS	.6406	.076638	.011619	.000000	.037257	13.04836	1.97825	14.15885	14.11524
URBAN	.7396	.195573	.010842	.059750	.010770	5.11318	.28346	10.23511	8.86134
RURAL	.2604	.123338	.010447	.032803	.018364	8.10780	.68675	12.11446	11.43271
NERGN	.2060	.126897	.008933	.041129	.014662	7.88041	.55475	12.98458	11.93929
MWRGN	.2450	.158612	.008640	.049784	.011003	6.30469	.34343	11.31364	10.14440
SRGN	.3448	.192451	.012613	.060829	.012635	5.19613	.34055	10.23667	8.90968
WRGN	.2042	.223828	.019007	.058919	.016804	4.46772	.37939	8.63056	7.56025
FARMS	.0180	.053637	.008885	.050833	.031801	18.64385	3.08836	19.87568	19.83747

3.2 Fitting Procedure and Empirical Results for ALLHSES

In this subsection, we estimate the equilibrium cumulative distribution function of the age of the process for *a randomly selected household.* Table 3 presents the fraction of households in each group (excluding FARMS) for $k = 1, \ldots, 8$. By using Equations (10) and (11) and the fitted cumulative distribution functions for the current residence time for the subgroups (OWNERS & RENTERS; URBAN & RURAL; or NERGN, MWRGN, SRGN, and WRGN), we obtain estimators for the cumulative distribution of the equilibrium current residence time as well as for the mean equilibrium current residence time for a randomly selected household, which equals $\hat{A}_0 = 10.7$ years. In Table 4 we provide three estimators for the total residence time of a randomly selected household whose membership is unknown, according to the three different partitions of the population. Similarly, we calculate three values of estimators, one for each partition, for $\hat{A}_0$; as we see in Table 4, $\hat{A}_0$ is not sensitive to the partition used. The figure for $\hat{A}_0$ under ALLHSES is obtained by Equation (11). Moreover, the estimator for the mean total residence duration of a randomly selected household is given by $\hat{\mu}_0 = 5.5$ years [see Eq. (13)].

3.3 Comparative Results

The average total residence duration calculated for all U.S. households, 5.5 years, is about half the average current residence time, 10.7 years. This result, which initially may seem counterintuitive, is caused, of course, by the greater weighing of short residence durations when calculating the average total residence duration than when calculating the average current residence time. In other words, in averaging total residence durations over a time interval, frequent movers may appear several times, whereas in averaging age residence times, each household appears only once. Indeed, Ioannides (1987), for example, used a panel of 1,151 families to investigate residential mobility and duration. The average age residence time for a span of 11 years was used to calculate time between moves of 106.87 months (8.9 years). The data for renters and owners were 39.34 months (3.28 years) and 185.86 months (15.5 years), respectively, compared to 2.45 years for renters and 13.0 years for owners given by our approach. On the other hand, Hempel and Ayal (1977) based their estimates on a rough residential mobility index indicating that less than 18% of Americans move in any specific year. This led to an estimate of about 6 years for average residence duration.

Based on the National American Housing Survey for the years 1974 to 1983, Morrow-Jones (1988) produced an average length of age residency of over 10 years as a base to their residence-duration estimate. Still a different approach and different results were shown by Rossi (1980). The data were derived from a household survey showing that during the entire life span the average number of addresses occupied by households was 3.16, leading to an estimate for average total residence time of about 10 years. In addition, the Rossi (1980) study predicted that 12% of the households will stay at the same residence for their entire expected life span, compared to less than 1% in our analysis.

Table 4. Values of $\hat{\mu}_0$, $s(\hat{\mu}_0)$, $\hat{A}_0$, and $s(T_0)$ Calculated for all Households From the Different Partitions of all Households

Partition	$\hat{\mu}_0$	$s(\hat{\mu}_0)$	$\hat{A}_0$	$s(T_0)$
RENTERS & OWNERS	5.106	.209	10.722	4.668
URBAN & RURAL	5.657	.271	10.724	7.543
NE & MW & S & W	5.648	.206	10.739	4.798

Hornik and Narayana (1982) also relied on respondents' willingness and ability to provide accurate data on their residency duration. Their survey, among others, attempted to use the data to estimate attrition rates of a panel. Consumers' responses produced an average age residence time of eight years. This statistic was used in their replacement and replenishment formulas of panel members.

Our data revealed that the expected total residence time varies notably across housing categories. It is about 2.4 years for renters (about 36% of all households) and 13 years for owners (about 64% of all households). Smaller, although significant, differences were also found among the different regions (from 7.9 years in the Northeast to about 4.5 years in the West) and between urban and rural areas—5.1 years (about 74% of all households) and 8.1 years (about 26% of all households), respectively.

4. SUMMARY AND CONCLUSIONS

The housing market is a major source of economic volatility. Therefore, reliable statistics are necessary to predict various economic trends. One such statistic is the age and the expected occupancy duration for all households and for different housing categories. These measures are important to explain many housing-related products and services, such as households' propensity to retain or replace furnishings and equipment, the buying of carpets, the investment in aesthetic home improvements, the maintenance of heating/cooling systems, and the purchasing of smoke alarms and home security systems and other consumer durables.

The common methods of estimating households' residence time vary and are of questionable validity. This article introduces a statistical procedure to estimate residence duration from census reports that provide extensive data on the age residence time. Specifically, by modeling the household's moving process, the residence-duration distribution is estimated from the age residence-time data. Using 1985, 1987, 1989, 1991, and 1993 U.S. housing survey data, distributions and averages for both the age and the total residence times are calculated for several housing categories. We found that the average total residence time for all U.S. households, 5.5 years, is less than half the average age residence time, 10.7 years.

The measurement techniques introduced in this article have several additional advantages over other commonly used measures. First, by using an intertemporal method of extracting information on duration of residence based on census data, the approach avoids the potential biases of direct-questioning approaches based on criteria supplied by the researcher. Second, the richness of the database used for calculating the desired statistics provided further insights into the nature of the time of residence for various housing categories and for different geographical regions.

The model and the calculations presented in this article predict the expected time for a *household* to stay in the same residence. (a) They do not predict the expected residence time for each member of the household, which is generally expected to be smaller. The values calculated here can be considered to represent upper limits of the expected time for individuals to live at the same residence. These values are, however, a more realistic estimate of the individual total residence time than is the average time a household has been living at its current residence. To calculate a more demographically dependent expected total residence time for individuals, the housing survey would have to include demographic data and residence duration data for each household member. Therefore, to provide further information for decision makers, future surveys should present data cross-tabulated against major consumer demographics and personal characteristics (e.g., by sex, age, income, minorities) as well as by type of housing units. (b) The model described in this article allows for intermove of households between segments. One approach to estimating the probability distribution of the total residence time of a household in a given segment, or alternatively of a randomly selected household, is by estimating the transition probabilities between segments and formulating the household moving behavior as a Markov switching model.

ACKNOWLEDGMENTS

We are grateful to Isaac Meilijson from the School of Mathematical Sciences at Tel-Aviv University for extensive discussions about the application of renewal theory and the derivation of the estimators. We also wish to thank the anonymous referee for his/her valuable suggestions and comments that greatly improved the article.

[Received February 1998. Revised July 1998.]

REFERENCE

Boehm, T. P. (1993), "Income, Wealth Accumulation and First-Time Homeownership: An Intermediate Analysis," *Journal of Housing Economics*, 3, 16–30.

Case, K. E. (1991), "The Real Estate Cycle and the Economy," *New England Economic Review*, 23, 38–46.

Crone, T. M., and Mills, L. O. (1994), "Forecasting Trends in the Housing Stock Using Age-Specific Demographic Projections," *Journal of Housing Research*, 2, 1–16.

Gerchak, Y. (1984), "Decreasing Failure Rates and Related Issues in the Social Sciences," *Operations Research*, 32, 537–546.

Golany, B., Phillips, F. Y., and Rousseau, T. T. (1991), "Few Waves vs. Continuous Consumer Panels: Some Issues of Attrition, Bias and Variance," *International Journal of Research in Marketing*, 8, 273–280.

Feller, W. (1966), *An Introduction to Probability Theory and Its Applications* (Vol. 2, 2nd ed.), New York: Wiley.

Harsman, B., and Quigley, J. M. (1991), *Housing Markets and Housing Institutions: An International Comparison*, Boston: Kluwer.

Hempel, D. J., and Ayal, I. (1977), "Transition Rates and Consumption Systems: A Conceptual Framework for Analyzing Buyer Behavior in Housing Markets," in *Consumer and Industrial Buyer Behavior*, eds. A. G. Woodside, J. N. Sheth, and P. D. Benett, New York: Elsevier, pp. 201–218.

Hornik, J., and Feldman, L. (1982), "An Empirical Investigation on the Do-It-Yourself Consumer Movement," *Journal of Retailing*, 58, 44–63.

Hornik, J., and Narayana, C. (1982), "Path-Type Reliability Assessments of Consumer Panel Data," *Journal of Economics and Business*, 34, 1–12.

Ioannides, Y. M. (1987), "Residential Mobility and Housing Tenure Choice," *Regional Science and Urban Economics*, 17, 265–287.

Israeli, M., and Nelson, C. B. (1992), "Distribution and Expected Time of Residence for U.S. Households," *Risk Analysis*, 12, 65–72.

Kidd, P. E. (1977), "Housing Statistics: Some Sources and Some Limitations," *American Real Estate and Urban Economics Journal*, 5, 335–354.

Landale, N., and Guest, A. (1985), "Constraints, Satisfaction and Residential Mobility," *Demography*, 22, 199–221.

Long, L. (1988), *Migration and Residential Mobility in the United States*, New York: Russel Sage.

Morrow-Jones, H. A. (1988), "The Housing Life-Cycle and the Transition From Renting to Owning a Home in the United States," *Environment and Planning*, 20, 1165–1184.

Nijkamp, P., Vanwissen, L., and Rima, R. (1993), "A Household Lifecycle Model for Residential Relocation Behavior," *Socio-Economic Planning Sciences*, 27, 35–54.

Pearson, R. W., Ross, M., and Dawes, R. M. (1992), "Personal Recall and the Limits of Retrospective Questions in Surveys," in *Questions About Questions*, ed. J. M. Tanner, New York: Sage Publications, pp. 65–94.

Pickles, A. R., and Davis, R. B. (1991), "An Empirical Analysis of Housing Careers: A Review and a General Statistical Modeling Framework," *Environment and Planning*, 23, 465–484.

Read, C. (1991), "Maintenance, Housing, Quality, and Vacancies Under Imperfect Information," *American Real Estate and Urban Economics Journal*, 19, 138–149.

Reinback, E. R., and Oliva, T. A. (1981), "General Living Systems Theory and Marketing: A Framework for Analysis," *Journal of Marketing*, 45, 30–37.

Ross, S. M. (1983), *Stochastic Processes*, New York: Wiley.

Rossi, P. H. (1980), *Why Families Move*, Beverly Hills, CA: Sage Publications.

Sharot, T. (1991), "Attrition and Rotation in Panel Surveys," *The Statistician*, 40, 325–331.

Shear, W. B. (1983), "A Note on Occupancy Turnover in Rental Housing Units," *American Real Estate and Urban Economics Journal*, 1, 525–532.

Soberon-Ferrer, H., and Dardis, R. (1991), "Determinants of Household Expenditures for Services," *Journal of Consumer Research*, 17, 385–397.

Sudman, S., and Ferber, R. (1977), *Consumer Panels*, Chicago: American Marketing Association.

U.S. Department of Commerce, Bureau of Census, and U.S. Department of Housing and Urban Development (1988), *American Housing Survey for the United States 1985*, Current Housing Reports (H-150-85), Washington, DC: Author.

——— (1989), *American Housing Survey for the United States 1987*, Current Housing Reports (H-150-87), Washington, DC: Author.

——— (1991), *American Housing Survey for the United States 1989*, Current Housing Reports (H-150-89), Washington, DC: Author.

——— (1993), *American Housing Survey for the United States 1991*, Current Housing Reports (H-150-91), Washington, DC: Author.