The Impact of Demographic Lifecycle States on Time to Vehicle Purchase: Insights from the Panel Study of Income Dynamics

Mohammad Mehdi Oshanreh
Department of Civil and Environmental Engineering
University of Washington, Seattle, Washington, 98195
Email: oshanreh@uw.edu

Nazmul Arefin Khan
Vehicle & Mobility Systems Department
Argonne National Laboratory, Lemont, Illinois, 60439
Email: khann@anl.gov

Don MacKenzie
Department of Civil and Environmental Engineering
University of Washington, Seattle, Washington, 98195
Email: dwhm@uw.edu

ABSTRACT

This study explores the impact of demographic lifecycle stages on the timing of car purchases, using data from the Panel Study of Income Dynamics from 1999 to 2021. Survival analysis was employed to model the duration until households purchase vehicles, incorporating key lifecycle variables such as age, employment status, marital status, childbirth, home ownership, and the presence of school-going children. The life table results indicate that early adulthood (ages 20-35) is the prime period for vehicle acquisition, with significant peaks around ages 25-30. Additionally, the instantaneous hazard of purchasing a vehicle is highest in the late 40s and early 50s. According to the Cox proportional hazards model, employment, marital status, and home ownership significantly increase the likelihood of purchasing a vehicle, while living in multi-unit dwellings decreases it. Interaction effects reveal that married individuals with employed spouses are substantially more likely to purchase vehicles. This study serves as a stepping stone in integrating demographic lifecycle analysis into car ownership modeling, making the modeling more reflective of real-world scenarios and improving the accuracy of policy and strategic planning.

Keywords: Car ownership, lifecycle states, survival analysis, panel study of income dynamics, PSID, demographic factors.

INTRODUCTION

Car ownership strongly influences the travel choices of individuals and households, affecting both short-term and long-term decisions. In the short term, owning a car can change daily activities and travel patterns, such as the types of activities people participate in, their destinations, modes of travel, and the availability of these modes. This, in turn, affects person miles traveled (PMT) and vehicle miles traveled (VMT) (*1–3*). In the long term, decisions about car ownership can influence where people decide to live and work (*3*, *4*). Despite the long history of car ownership modeling in the literature, limited research exists on how shifts in household composition, employment changes, and other lifecycle transitions impact household car ownership decisions and trends. This study aims to fill this gap by using panel data that captures household transitions and changes over time.

Since the 1930s, car ownership models have been developed for various purposes, such as evaluating consumer preferences, predicting future demand, guiding investment decisions, estimating tax revenues, assessing regulatory impacts, and forecasting transportation demand, energy use, and emissions. These models are used by car manufacturers, oil companies, international organizations, and different levels of government (*5*).

In the past, car ownership models mostly relied on aggregate time series data. One common method used was a sigmoid-shaped function to represent the growth of car ownership over time (*5*). This growth was often linked to economic indicators like income levels or GDP per capita (*6*). For example, Romilly et al. (*7*) created a regression-based model to predict per capita car ownership, using variables like personal disposable income per capita, motoring cost index, bus fare index, unemployment rate, road network size, interest rate, and age structure. Similarly, Dargay and Gately (*8*) used a time series model to study the effect of income on car ownership. However, these aggregate models fall short in capturing the behavioral dynamics that drive household decision-making (*9*), thus limiting their accuracy and effectiveness in real-world policy applications.

More recently, research on car ownership has shifted to disaggregate models, which focus on individual households or people rather than the entire population. These models consider socio-economic factors, geographical locations, ownership costs, and the availability of alternative transportation options, providing a more detailed understanding of the factors that influence car ownership decisions (*5*, *10*). In the early 1990s, De Jong (*11*) proposed a microeconomic utility model that included fixed and variable car costs within the budget constraint. Using revealed preference (RP) survey data, he found that both fixed and variable car costs effectively reduced VMT. Fixed costs mainly affected car ownership levels and indirectly influenced VMT, while variable costs had a direct impact on car usage.

Linciano (*12*) developed a regression model linking mileage driven to demographic variables (e.g., age and income), locational attributes (e.g., transit availability), and vehicle characteristics (e.g., engine size, vehicle age). Brownstone et al. (*13*) used multinomial logit (MNL) and mixed logit (MXL) models to analyze consumer preferences for alternative fuel vehicles, considering fuel type, vehicle range, price, home refueling cost and time, service refueling time and cost, acceleration, top speed, tailpipe emissions, vehicle size, and body type. Choo and Mokhtarian (*14*) applied an MNL model to examine the influence of various factors on vehicle type choice, including both vehicle characteristics and demographic factors.

Although most prior work has taken advantage of disaggregate data in modeling individual car ownership decisions, a key limitation is that they generally use only cross-sectional data. This means they capture a snapshot of a specific point in time and fail to account for the dynamic changes in demographic and economic factors over a person's lifecycle. As a result, these models often overlook the time-varying aspects of behavior and context, such as changes in income, family size, employment status, and housing location. This limitation can affect the accuracy of predictions and the ability to develop effective long-term policies and strategies.

Survival analysis, a statistical method initially developed for medical and engineering fields to study time-to-event data (*15–18*), has since been widely adopted across various disciplines, including transportation research. This technique models the duration until a specific event occurs, such as the adoption of a new transportation mode, vehicle replacement, or the occurrence of traffic crashes.

In the realm of transportation safety, survival analysis has been employed to study the duration until traffic crashes occur. This application considers various factors, including driver behavior, road conditions, and vehicle characteristics, to identify high-risk elements and develop targeted safety interventions. For instance, Shankar et al. (*19*) used survival models to assess the impact of roadway geometrics and environmental factors on rural freeway crash frequencies. Similarly, Abdel-Aty and Radwan (*20*) applied survival analysis to model traffic crash occurrence and involvement, considering factors such as weather conditions, time of day, and driver demographics.

Survival analysis is also instrumental in studying travel behavior patterns over time. Researchers use this method to understand how travel habits evolve and how different policy measures influence these behaviors. Juan and Xianyu (*21*) applied survival analysis techniques to evaluate travel time costs, linking travel demand strategies and transportation control measures to time. This study uses a duration model to analyze commuters' daily travel time, examining its relationship with sociodemographic factors and activity patterns.

Furthermore, researchers have used survival models to investigate the time to adopt new transportation modes by individuals and households, such as public transit options. Bhat and Steed (*22*) used a continuous-time model to study the departure time choices for urban shopping trips, illustrating how socio-economic characteristics and geographic factors influence the adoption timing of new transportation modes.

Beyond travel behaviors and safety, survival analysis is widely used to model household vehicle transactions and car ownership dynamics. Mohammadian and Rashidi (*23*) employed a competing risk duration approach to understand how factors such as leasing, vehicle age, and economic conditions influence transaction timing. Khan and Habib (*24*) expanded on this by analyzing the timing of vehicle transactions and the types of vehicles chosen, using a survival analysis framework. Ghasri et al. (*25*) compared survival analysis with discrete choice models to simulate vehicle ownership dynamics, particularly focusing on vehicle transactions. Similarly, Mannering and Winston (*26*) explored the factors influencing the duration of car ownership and the timing of vehicle replacement decisions, incorporating variables such as household income, fuel prices, technological advancements, and policy interventions. Ramjerdi et al. (*27*) highlighted the use of survival analysis in modeling car ownership, transactions, and vehicle type. Yamamoto et al. (*28*) employed a competing-risk duration model to model vehicle transactions. Gilbert (*29*) developed a hazard model to estimate the distribution of vehicle ownership lengths.

To summarize, various disaggregated models have been developed to analyze household and individual vehicle ownership decisions. Literature largely relies on cross-sectional data due to its wider availability and ease of collection. However, there have been efforts to use panel data to track individuals over time, aiming to understand how changes in household dynamics affect vehicle ownership choices. For instance, Mohammadian and Miller (*30*) used a 9-year panel dataset to connect household characteristics—such as members' educational attainment, average age, and driver attributes, along with fleet details like fleet age and vehicle class—to different purchasing scenarios (new, secondhand, used, old). Similarly, Jensen et al. (*31*) used a stated preference long-panel survey to examine the choice between conventional gasoline/diesel cars and electric vehicles, finding that various vehicle attributes (e.g., purchase price, driving cost, range, and charging infrastructure availability) significantly impact the decision between conventional and electric vehicles.

Although duration models have a long history in transportation research, including the study of vehicle transaction timing, most prior work overlooks how key lifecycle states—such as employment status, age, childbirth, homeownership, and housing type (single-family or multi-unit dwelling)—impact the timing of changes in car ownership. This gap exists largely due to the limited availability of panel data that capture these aspects and transitions alongside transportation-related decisions. By using panel data to track household transitions over time, this study aims to provide a more comprehensive view of how these life stages influence vehicle purchase timing.

DATA

The Panel Study of Income Dynamics (PSID) is the world's longest ongoing longitudinal survey of households. Initiated in 1968, it started with a nationally representative group comprising more than 18,000 people in 5,000 families across the United States (*32*). Continual collection of data on these individuals and their offspring encompasses a wide array of topics, such as employment, income, wealth, spending habits, health, marital status, childbirth, child development, charitable giving, education, and much more. The University of Michigan manages the PSID. Certain categories of data within the PSID are accessible solely through a restricted data use agreement. This restricted data encompasses, among others, geospatial information more detailed than state level, mortality records, data on assisted housing provided by the U.S. Department of Housing and Urban Development, Medicare claims, and information regarding educational traits sourced from the National Center for Education Statistics.

A complete description of more than 6,000 variables can be found on the PSID website. Most of the information is gathered at the household level. Below, we will briefly discuss the key lifecycle variables used for vehicle purchase modeling. The PSID data were collected annually from 1968 to 1997 and biennially after 1997. Some variables were gathered at specific times and may not have been consistently collected throughout the entire study duration. Table 1 shows the variables used in this study and their corresponding coverage. An important detail is that in the survey file structure, the reference person is typically the main male in the household. If the reference person passes away or is not present in the household and there is no other eligible male, the spouse becomes the reference person. More details on this eligibility can be found on PSID documentation (*32*).

TABLE 1 PSID Variables Used in this Study

Section	Information	Coverage
Children	The age of the youngest child under 18 in the family unit (FU) (children at home only).	1968-2021
Demographic	Age of the reference person and spouse.	1968-2021
Demographic	Marital status of the reference person.	1977-2021
Housing	Own or rent status of the current dwelling.	1968-2021
Housing	Type of dwelling structure (single-family or mutli-unit dwelling (MUD))	1968-1972 1975-1981 1983-2021

Section	**Information**	**Coverage**
Retirement	The actual year in which reference person or spouse retired.	Reference person: 1968-2021. Spouse: 1978-2021
Work	Employment status of reference person and spouse.	1968-2021
Expenditures	Number of cars owned by a household.	1968-1972 1975-1986 1999-2021
Expenditures	If the second, third, and most recent vehicles are purchased within the past two years leading up to that survey wave.	1999-2021

Due to limited funding and a significant increase in the original 1968 sample of families from split-off families, approximately 2,500 families were removed from the PSID sample in 1996, with this reduction taking effect from 1997 onward. Additionally, information about transportation expenditures, particularly vehicle acquisition, has been incorporated into the survey questionnaire since 1999. We have used publicly available data from the 1999-2021 waves in this study.

The publicly available vehicle-related questions focus on the number of vehicles a household owns and whether they have acquired a vehicle within the two years leading up to the survey wave. Information is collected only for the three most recent vehicles, sorted by age. Details about the make and model of these vehicles are accessible only through a restricted dataset and are provided under a restricted-use contract between the researcher and the University of Michigan. Since the survey is conducted biennially, our data is heavily interval-censored, meaning the exact timing of a household's car purchase is unknown. After the data cleaning process, information about 17,000 households was used in the modeling.

Descriptive Analysis of the Data

In this subsection, we briefly describe the variables used directly and indirectly in this study. The original format of the data is wide, which means that all the information for each household is recorded in a single row. This format requires an extensive data cleaning process, as the same variable is coded differently for each specific wave. An example of the relabeled data, including a few variables for four households is shown in Table 2.

Based on Table 2, each household has one row of data spanning from 1968 to 2021 (with 1999 and 2001 shown here as examples). For the years in which information was gathered for a specific household, the columns contain values; otherwise, they are marked as NA. Row 1 represents an individual who was not married in 1999 but got married sometime between 1999 and 2001 (before the 2001 survey). This demonstrates how the data is interval-censored, meaning the exact date of a demographic change is unknown. Row 2 represents an individual who was present in 1999 but became right-censored in 2001, either due to death or being lost to follow-up. Row 3 shows an individual who was the reference person in 1999 but got married between 1999 and 2001. Subsequently, they became the spouse, as a non-sample male joined the household and became the new reference person. Row 4 shows an individual who was unemployed and unmarried in 1999 and was still unmarried in 2001 but had become employed.

TABLE 2 An Example of PSID Raw Data with relabeled columns. Only a few rows and columns are shown, out of the hundreds of columns and thousands of rows in the data set. Each row represents a unique household.

Row	1999 Family ID [ER13002]	1999 Reference person Age [ER13010]	1999 Spouse Age [ER13012]	1999 Employment Status of Reference Person [ER13205]	2001 Family ID [ER17002]	2001 Reference person Age [ER17013]	2001 Spouse Age [ER17015]	2001 Employment Status of Reference Person [ER17216]
1	4309	75	0 (No Spouse)	4 (Retired)	4002	77	65	4 (Retired)
2	3.0	76	0 (No Spouse)	4 (Retired)	NA	NA	NA	NA
3	6965	20	0 (No Spouse)	1 (Working now)	7308.0	31	22	1 (Working now)
4	5326	20	0 (No Spouse)	3 (Looking for work, unemployed)	5367	22	0 (No Spouse)	1 (Working now)

Notes: Column name entries in brackets are the original variable names in the PSID raw data file. Table entries in parentheses are decoded values from the PSID codebook.

Below, we present a summary of some key variables. Figures 1 and 2 show the average and median number of vehicles per household by year, and the proportion of employment statuses of the reference person by age group, respectively.

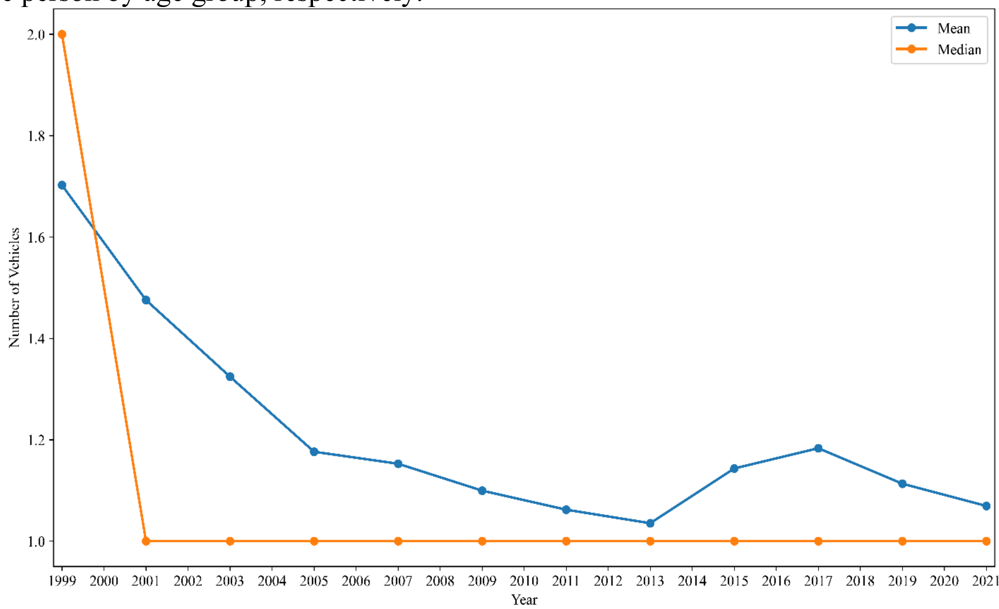

Figure 1 Mean and Median number of vehicles per Household by Year

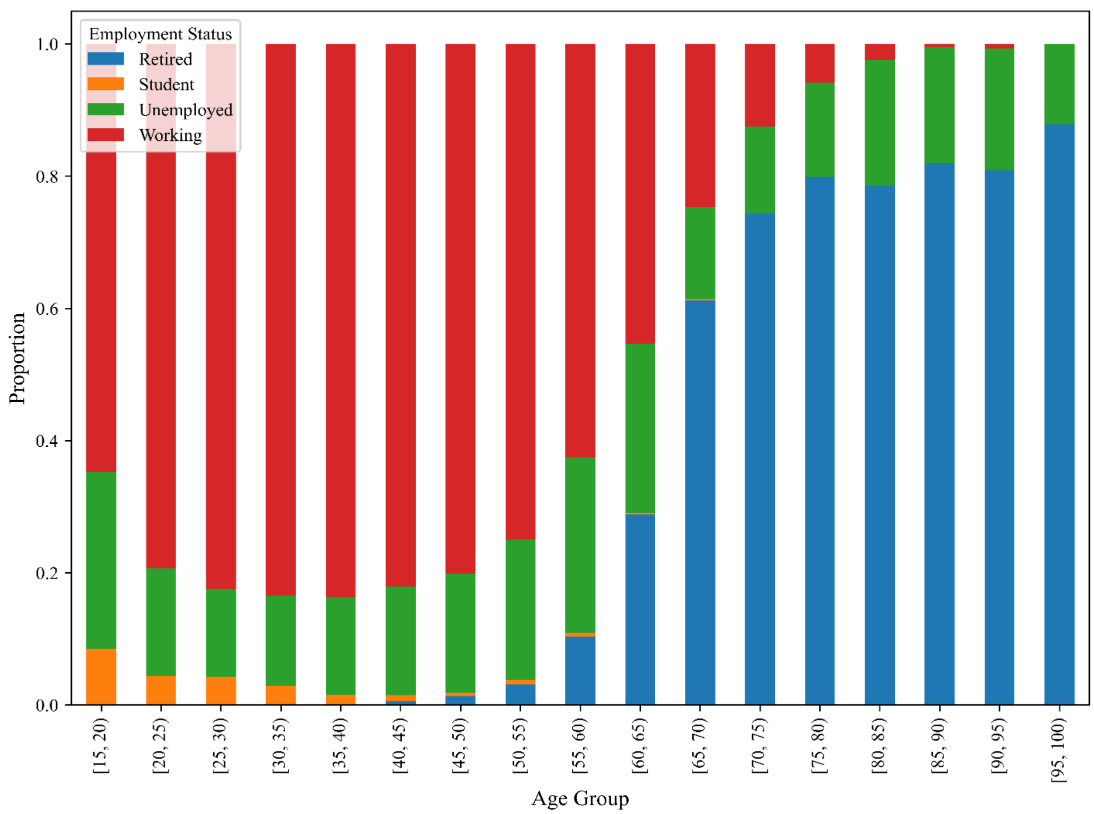

Figure 2 Employment Status of the Reference Person by Age Group

To demonstrate the representativeness of the PSID, we compared key demographic variables to the American Community Survey (ACS) 1-year estimates, except for 2020, where the 1-year estimate was unavailable, and the 5-year estimate was used. Figure 3 illustrates the median age of individuals in the PSID data by gender and year and the corresponding ACS estimates of the median age of the U.S. population. However, we did not conduct any formal statistical tests to verify this, as these validations are available in the PSID documentation.

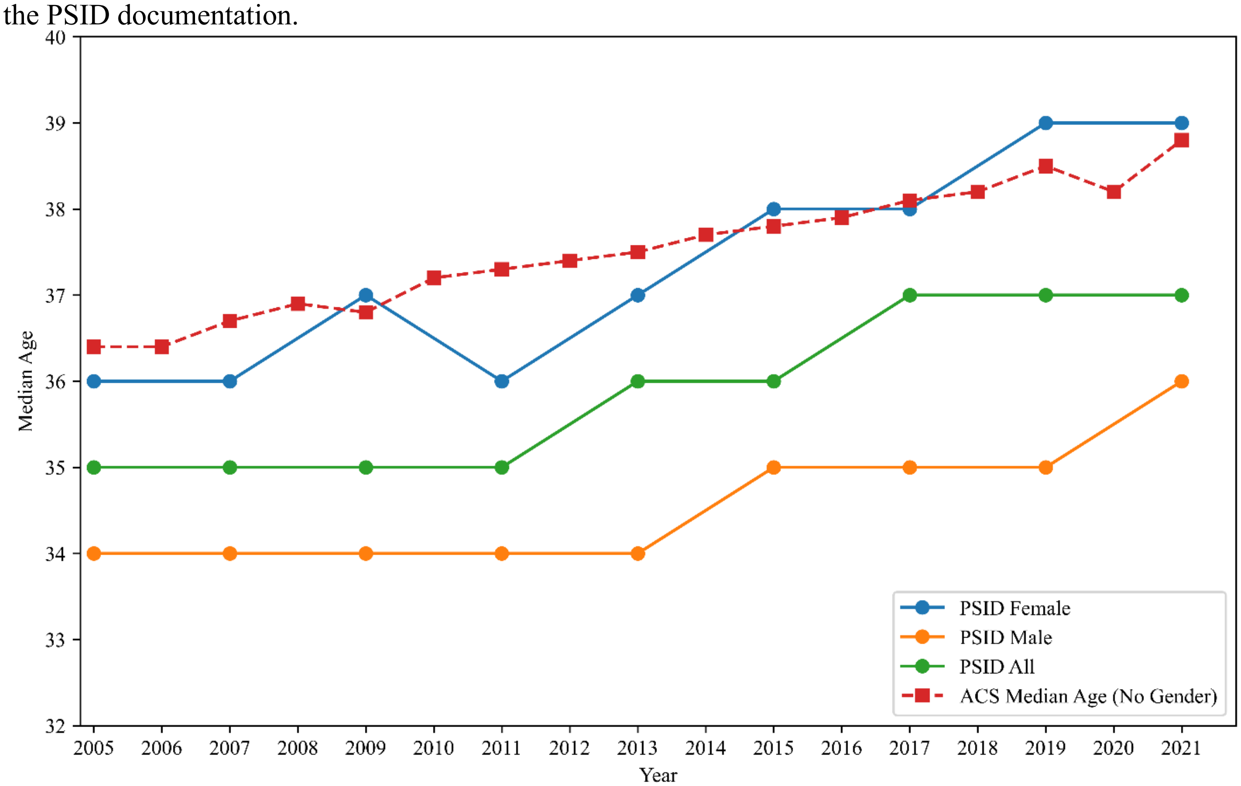

Figure 3 Comparison of PSID and ACS Median Age by Year

Figure 4 shows the proportion of male individuals in each wave of survey and its corresponding value in the ACS.

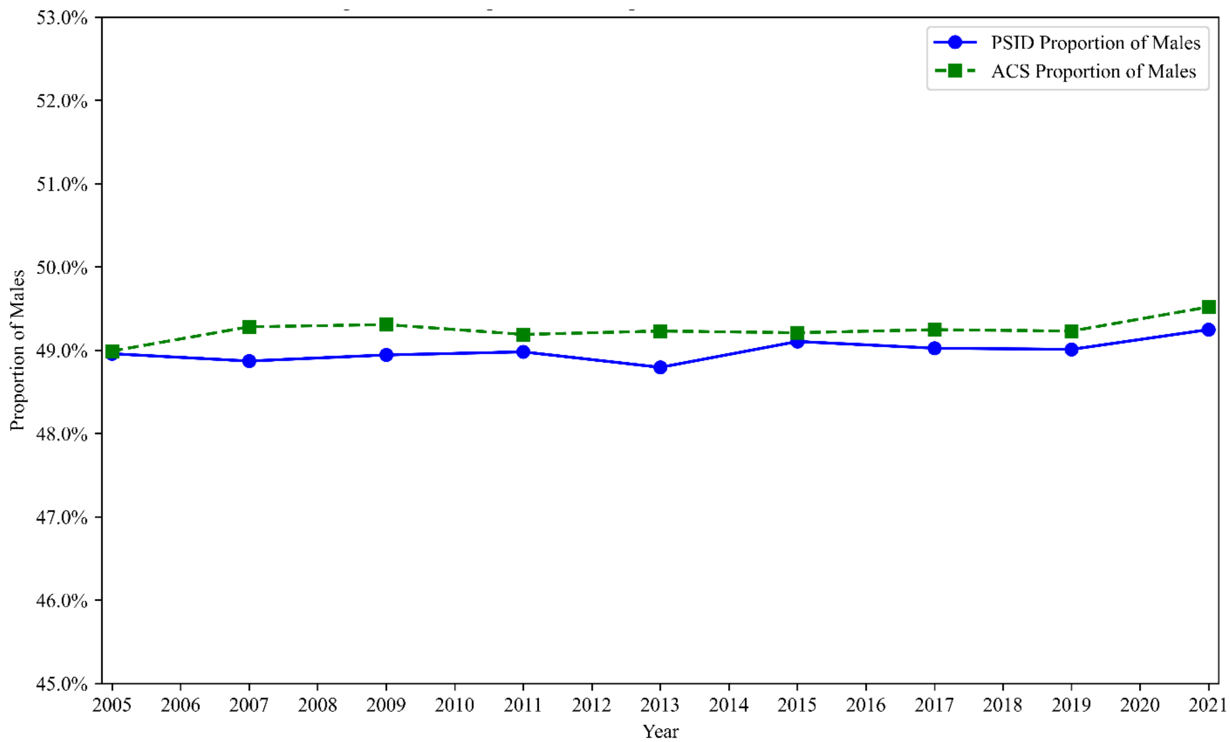

Figure 4 Comparison of PSID and ACS Proportion of Males

Figure 5 shows the median of total household income by year. The values are in nominal US dollars.

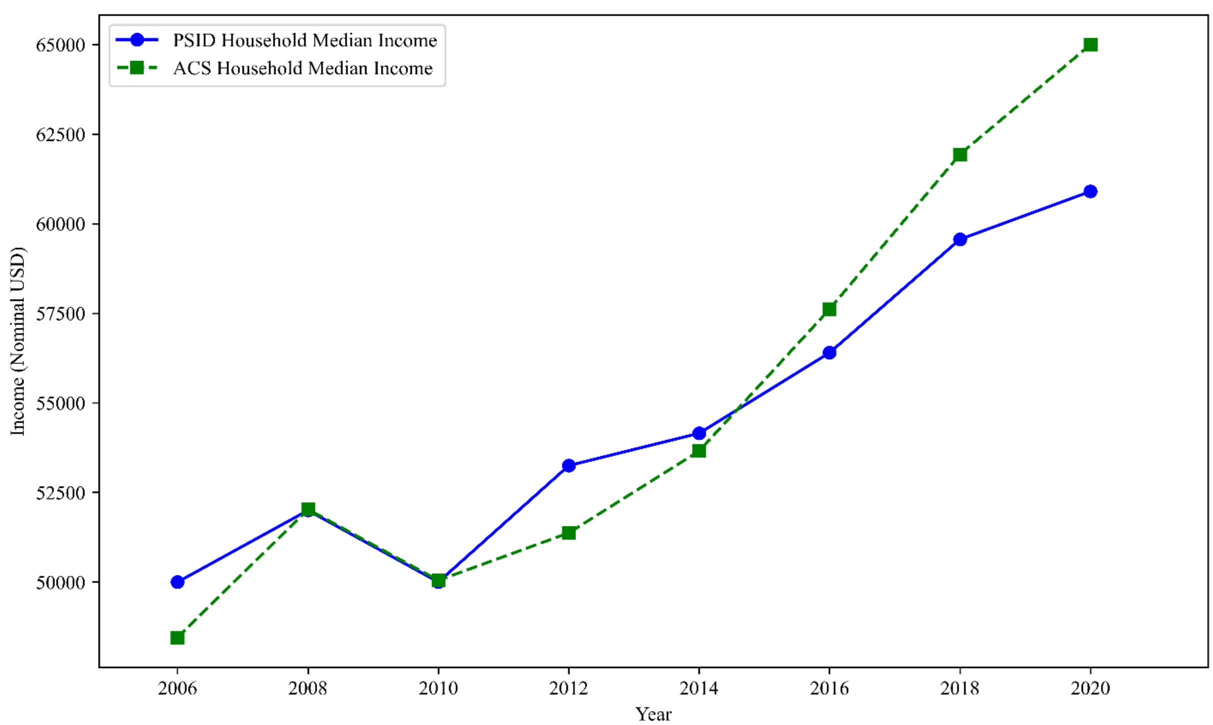

Figure 5 Comparison of PSID and ACS Median Household Income in nominal US dollars

To summarize, Figure 1 shows the median and mean number of vehicles per household, which generally decreases over time, except for a brief rise in the mid-2010s. Figure 2 presents the employment status of the reference person across age bins. As expected, the proportion of students declines with age, employment peaks in middle age, and retirement rates increase significantly from the 50s onward. Figure 3 demonstrates a good alignment between PSID age distribution and that of the overall US population, with both trending upward over time. Figure 4 compares the gender distribution in PSID to that of the US population, revealing a close match. Figure 5 compares the median household income in PSID with national data. The alignment remains strong through the late 2010s, with small discrepancies emerging thereafter, possibly attributable to the economic disruptions caused by COVID-19 and the reliance on ACS 5-year estimates rather than 1-year estimates.

Overall, the PSID sample closely mirrors the US population in terms of key demographic and socioeconomic characteristics.

METHODS

Survival analysis is a branch of statistics focused on time-to-event data. Its primary objective is to model and interpret the duration until specific events of interest occur, such as death, mechanical failure, or disease relapse. Within this field, there are two principal approaches: parametric and non-parametric survival models. The difference between these approaches lies in their assumptions about the underlying survival function. Parametric models assume a specific statistical distribution for the survival times. In contrast, non-parametric models do not assume any predefined distribution for the survival function. Moreover, when it comes to considering time-varying covariates such as age, family composition and employment status, non-parametric models are preferred (*33*). Since non-parametric models do not assume any specific form for the survival time distribution, they are more flexible when the underlying distribution is unknown or complex. Life tables, a non-parametric approach, are widely used in literature to analyze time-to-event data and gain insights into the underlying patterns of the outcome, especially in the presence of interval and right-censored data. (*34–36*).

Life tables

The Kaplan-Meier estimator (*37*) was employed to estimate the survival function from the data. This nonparametric statistic allows for the estimation of the survival function without assuming any particular form for the underlying hazard function. The Kaplan-Meier survival curve is a step function that changes value at each event time. To compute the Kaplan-Meier estimator, we let T_i represent the observed survival times, with d_i denoting the number of failure events (e.g., car purchases) at time T_i, and n_i representing the approximate number of individuals (e.g., households in this study) at risk just prior to time T_i. The estimator $S(t)$ of the survival function at time t is given by:

$$S(t) = \prod_{T_i \leq t} \left(1 - \frac{d_i}{n_i}\right) \tag{1}$$

where the product extends over all distinct times of observed events up to and including time t. To account for censoring, an equivalent number of individuals is estimated as shown in Eq. (2):

$$n_i = N_i - 0.5c_i \tag{2}$$

where N_i is the number of individuals surviving at the start of interval i and c_i is the number of individuals who have left the study during interval i without experiencing the failure event. The intuition behind the 0.5 is that we assume a uniform distribution of censoring at each interval (*38*).

The Nelson-Aalen estimator (*39*, *40*) was used to estimate the cumulative hazard function. This estimator provides a nonparametric approach to estimate the cumulative hazard rate based on observed survival data, similar to the Kaplan-Meier method. The cumulative hazard function $H(t)$ at time t is estimated using:

$$H(t) = \sum_{T_i \le t} \frac{d_i}{n_i} \tag{3}$$

where T_i, d_i and n_i are defined earlier.

The conditional probability of failure (q_i) is computed as the proportion that failed in the interval, given that an individual has survived until that interval. p_i is the conditional proportion surviving that interval. These are computed as shown in Eq. (4):

$$q_i = \frac{d_i}{n_i}; \quad p_i = 1 - q_i \tag{4}$$

The probability density function (PDF) can be calculated as shown in Eq. (11):

$$f(t) = \frac{S(t_i) - S(t_{i+1})}{w_i} \tag{5}$$

where w_i is the width of the ith interval and $S(t_i)$ is defined in Eq. (1).

Non-Parametric Survival Models with Covariates

The Cox proportional hazards model is the most used non-parametric method in survival analysis, allowing for the inclusion of covariates without specifying the baseline hazard function (*34*). The Cox model can be expressed as:

$$h(t|X) = h_0(t)\, exp(\beta' X) \tag{6}$$

where $h(X)$ is the hazards function at time *t* given the vector of covariates *X*. In this model, $h_0(t)$ is the unspecified baseline hazard function, and β' is a vector of regression coefficients. The partial likelihood approach is used to estimate the regression coefficients. The partial likelihood function for n individuals is:

$$L(\beta) = \prod_{i=1}^{n} \left(\frac{exp(\beta' X_i)}{\sum_{j \in R(t_i)} exp(\beta' X_j)} \right) \tag{7}$$

where $R(t_i)$ is the risk set at time t_i, the set of individuals still at risk just before time t_i. The log-partial likelihood function is then maximized to obtain the estimates of β.

To summarize, there are multiple methods to model the timing of vehicle purchases. Parametric models are appropriate when the underlying distribution is known or can be estimated, and when the covariates, such as sex, remain constant. Non-parametric models offer greater flexibility as they do not require assumptions about the underlying distribution. Life tables are valuable for understanding the overall phenomenon of interest but do not account for covariates. For this reason, we later used a Cox proportional hazards model to assess the specific effects of lifecycle covariates on the timing of car ownership. Since a starting point for risk must be chosen, we selected age 15 of the head of the household, as this is the minimum age at which an individual can obtain a leaner's driving permit in most states in the USA. We first constructed a life table to broadly assess the distribution of the timing of vehicle purchases. Following this, we fitted a Cox proportional hazards model to the data to examine the factors influencing the timing more closely.

RESULTS and DISCUSSION

Table 2 displays the life table for the timing of vehicle purchases. N_i indicates the number of households who are at risk of purchasing a vehicle at interval i, d_i s the number of households who have purchased a vehicle (failed) at interval i, and c_i is the number of households who were right-censored (lost to follow-up) at that interval. Figure 6 shows the empirical PDF calculated from the life table. The results suggest that the likelihood of purchasing a vehicle is highest between the ages of 20 and 35, with a noticeable peak in the 25-30 age range, suggesting that early adulthood is the prime period for acquiring a vehicle. This trend gradually declines as individuals age, with significantly fewer purchases occurring in later life stages, particularly after age 60. These patterns highlight early adulthood as the most critical demographic for vehicle purchases and suggest that policies aimed at curbing car ownership should primarily target individuals in this age group.

According to Figure 7, the hazard function peaks in the late 40s and early 50s, indicating that the conditional instantaneous risk of buying a car is highest during this period. This suggests that, conditional on not having purchased a car up to that point, the risk of buying a car is maximized in this age range. This may be due to factors like career stability, increased income, preparation for retirement, or changing family needs as children approach independence. Additionally, there is a notable spike in the late 80s, but data in these intervals are heavily right-censored, which inflates the hazard function as the number of effective households decreases. As evident, the 95% confidence intervals are getting wider in older age groups as the ratio of censored individuals to at-risk individuals increases. However, unseen factors such as financial changes or retirement investment tendencies might also contribute to this peak.

TABLE 2 Life Table of Time to Purchase a Vehicle

Age group (*t*)	At Risk (N_i)	Purchased (d_i)	Censored (c_i)	Effective at risk (n_i)	Purchasing prob (q_i)	Surviving prob (p_i)	Survival function *S(t)*	PDF	Hazard function *(H(t))*
15 -20	17254	130	119	17194.5	0.008	0.993	1	0.002	0.002
20-25	17005	1809	902	16554	0.11	0.891	0.993	0.022	0.024
25-30	14294	2493	780	13904	0.18	0.821	0.884	0.032	0.04
30-35	11021	1913	445	10798.5	0.178	0.823	0.726	0.026	0.039
35-40	8663	1569	357	8484.5	0.185	0.816	0.597	0.023	0.041
40-45	6737	1337	254	6610	0.203	0.798	0.487	0.02	0.046
45-50	5146	1173	227	5032.5	0.234	0.767	0.389	0.019	0.053
50-55	3746	881	172	3660	0.241	0.76	0.298	0.015	0.055
55-60	2693	558	128	2629	0.213	0.788	0.227	0.01	0.048
60-65	2007	382	142	1936	0.198	0.803	0.179	0.008	0.044
65-70	1483	297	138	1414	0.211	0.79	0.143	0.007	0.047
70-75	1048	247	163	966.5	0.256	0.745	0.113	0.006	0.059
75-80	638	130	152	562	0.232	0.769	0.085	0.004	0.053
80-85	356	93	96	308	0.302	0.699	0.065	0.004	0.072
85-90	167	45	80	127	0.355	0.646	0.046	≈ 0	0.087

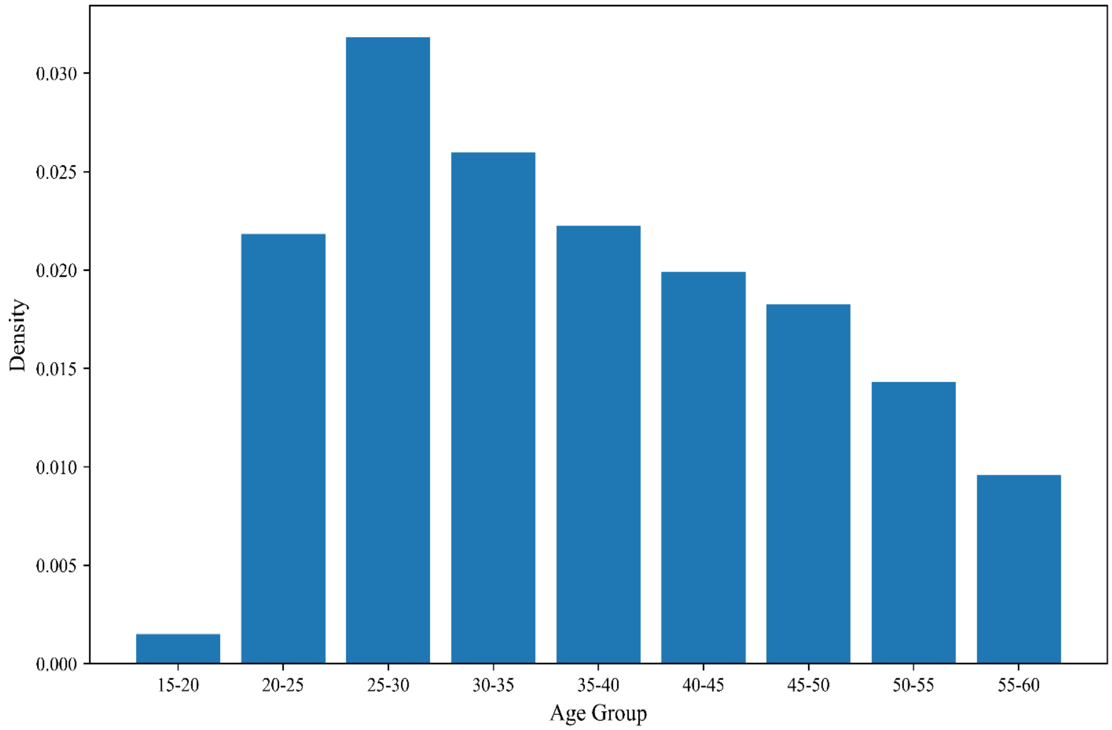

Figure 6 Empirical PDF of Age at Vehicle Purchase

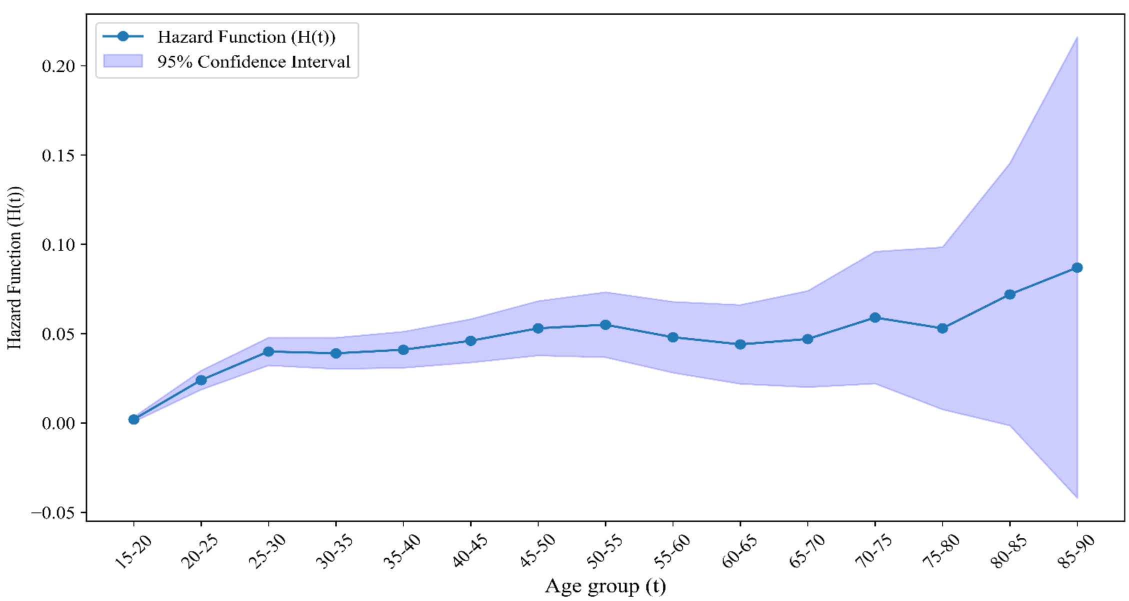

Figure 7 Hazard Ratio of Age at Vehicle Purchase

Table 3 presents the results of the Cox proportional hazards model. While most of the variable names are self-explanatory, some variables have specific baseline values that are important to understand. For instance, the baseline value for the employment status of the reference person is that they are not working, which includes individuals who are either not currently working or are students. The baseline marital status for the reference person is unmarried, encompassing those who are divorced, separated, or single. The baseline for home ownership indicates that the household is renting their current residence. The baseline for the housing structure is single-family housing. School-going children are those aged between 6 and 18 years. Additionally, we have included an interaction term between marital status and the employment status of the spouse. This interaction variable highlights the effect of the spouse's employment status when they are present in the household.

TABLE 3 Cox Proportional Hazards Model Results

Variable	coef	exp(coef)	se(coef)	*z*	exp(-coef)	P-value
If reference is employed	0.772	2.164	0.029	26.883	0.462	<0.001***
If reference is retired	0.506	1.659	0.054	9.448	0.603	<0.001***
If reference person is married	0.849	2.336	0.027	31.419	0.429	<0.001***
If lives in MUD	-0.315	0.729	0.023	-13.823	1.371	<0.001***
If owns the house	0.391	1.478	0.023	16.928	0.677	<0.001***
If child was born in the two years preceding the survey wave	0.0320	1.033	0.033	0.984	0.968	0.325
If School-going child(ren) present	0.109	1.115	0.023	4.652	0.897	<0.001***
If married * If spouse employed	0.176	1.193	0.026	6.647	0.838	<0.001***
n= 33811, number of events = 12998						
Concordance = 0.708 (se = 0.003)						
Likelihood ratio test = 7,077 on 8 df, p <0.001***						
Wald test = 6,740 on 8 df, p<0.001***						
Score (logrank) test = 7,707 on 8 df, p <0.001***						
AIC: 205052						
BIC: 205112						
Note: * significant at 0.1 level, ** significant at 0.05 level, *** significant at 0.01 level						

The results indicate employment status of the reference person has a significant impact on the likelihood of purchasing a vehicle. Employed individuals are approximately 2.16 times ($p < 0.001$) more likely to purchase a vehicle compared to those who are not currently working (but not retired). Our results align with the conclusions of earlier research (*41*). Similarly, retired individuals are about 1.66 times ($p < 0.001$) more likely to purchase a vehicle compared to those who are not currently working.

Marital status also plays a crucial role, with married individuals being approximately 2.33 times ($p < 0.001$) more likely to purchase a vehicle compared to unmarried individuals. The results we obtained are in agreement with earlier studies. A study by Oakil et al. (*42*) demonstrates the relationship between household formation and car ownership changes. Living in a multiple-unit dwelling (MUD) decreases the hazard of purchasing a vehicle by about 27% ($1 - e^{-0.315} \approx 0.27$), which is statistically significant (p <

0.001). In contrast, owning a home increases the likelihood of purchasing a vehicle by about 48% ($p < 0.001$). These observations validate prior findings in the literature (*43*).

The interaction term indicates that being married and having an employed spouse together increases the likelihood of purchasing a vehicle, with a hazard ratio of 1.19. This means that married couples with an employed spouse are approximately 1.19 times more likely to buy a vehicle compared to married individuals whose spouse is not employed. To understand the combined influence of being married and having an employed spouse, we add the individual effect of being married (0.849) and the interaction effect (0.176), resulting in a total coefficient of 1.025. The combined hazard ratio is $e^{1.025} \approx 2.78$. This indicates that married individuals with an employed spouse are approximately 2.78 times more likely to purchase a vehicle compared to unmarried households without an employed spouse.

Childbirth event appears to slightly increase the hazard of purchasing a vehicle, but this effect is not statistically significant ($p = 0.32$). This does not mean that childbirth has no effect on the hazard rate, since a childbirth would ultimately affect the presence of school-going children which increases the likelihood of purchasing a vehicle by about 11.5% ($p < 0.001$). Prior research showed the relationship between school escort trips and car ownership (*44*, *45*).

In summary, the results suggest life cycle states and events, such as employment status, marital status, home ownership, and household size, are good predictors of vehicle purchase timing. Employed and retired individuals, married individuals (especially with an employed spouse), and homeowners are more likely to buy a vehicle. The presence of school-going children also increases this likelihood. The model's concordance of 0.708 suggests strong predictive ability, and the highly significant likelihood ratio, Wald, and Score (logrank) tests ($p < 0.001$) confirm a good fit.

CONCLUSIONS

Car ownership profoundly shapes short-term travel behaviors, influencing daily activities, mode choices, and distances traveled, impacting both PMT (person-miles traveled) and VMT. The scarcity of longitudinal data on households' car ownership decisions limits the use of life history models in vehicle transaction modeling. Existing datasets often do not capture important lifecycle events at the household or individual level. Using data from the PSID, we aimed to answer how various lifecycle events influence the timeline for purchasing a vehicle.

The results of our study indicate that significant lifecycle states such as employment, marital status, and home ownership heavily influence car ownership decisions. Our findings align with previous research, such as Clark et al. (*46*), who demonstrated that life events like changes in employment status, residential relocation, and family composition significantly impact household car ownership decisions. Similarly, Beige and Axhausen (*47*) and Oakil et al. (*42*) highlight those transitions, such as moving out of the parental home and having a new child, lead to changes in car ownership patterns. These studies underscore the importance of considering life events in vehicle transaction modeling, as they play a crucial role in shaping car ownership decisions. Furthermore, research by Bhat and Koppelman (*48*) supports our results by showing that employment and income are key determinants of household car ownership within a dynamic travel demand framework.

Our results highlight life-cycle events as "windows of opportunity" for transportation planners and policymakers to influence car ownership decisions (*49*, *50*). These findings suggest that transportation planners can leverage key lifecycle stages—such as employment stability, marriage, and homeownership—to design policies that promote sustainable travel habits. For example, planners might introduce incentives for public transit or shared vehicles aimed at newly married couples or homeowners, where the likelihood of car ownership increases. Practical interventions could also include subsidizing cargo bikes or e-bikes for young families to delay car purchases, encouraging family-friendly housing in transit-rich areas, and investing in quality schools in these neighborhoods to prevent families from moving to car-dependent suburbs. Although certain life events, like household formation or dissolution, may be less amenable to direct policy interventions, planners can make alternative transportation modes more attractive and accessible, helping to reduce the dependence on car ownership across various life stages. Additionally, this framework can be applied in activity-based models to simulate the timing of vehicle purchases, enriching predictive accuracy in these models by integrating lifecycle-driven decision patterns.

There are several limitations to this work, which future research can address to enhance the modeling framework. A key challenge we faced was that our data is interval-censored, meaning the exact timing of vehicle purchases and lifecycle changes is not observed. Additionally, as individuals enter the survey, no information is collected about their prior years, limiting our understanding of their past experiences if they were not part of the survey from the beginning. Additionally, other vehicle transactions, such as trading in or shedding a vehicle, could be considered, but this information is not available in the PSID's public dataset. Using the restricted dataset this information can be inferred.

Another limitation is that car ownership information was gathered at the household level, preventing us from identifying individual users of the vehicles within the household. Incorporating interaction and competition over vehicles among household members could enhance the model and better reflect real-world scenarios. Since 2011, information about vehicle engine types (gasoline, hybrid, electric, etc.) has been included in the PSID. However, with only seven rounds of this data available, it is currently too early to include vehicle types in the analysis. As more data is gathered in future rounds, researchers will be able to incorporate this information into their analyses.

Moreover, prior literature suggests that the accessibility measures, built environment and the availability and characteristics of non-car transportation modes heavily influence car ownership decisions (*51–54*). Incorporating residential and employment locations of household individuals into the model can account for these factors. However, this approach necessitates careful consideration of ethical issues related to the use of individuals' home and work addresses.

ACKNOWLEDGMENTS
This work was supported by the U.S. Department of Energy (DOE) Vehicle Technologies Office (VTO) under the Transportation Systems and Mobility Tools Core Maintenance and Pathways to Net-Zero Regional Mobility, an initiative of the Energy Efficient Mobility Systems (EEMS) Program. The submitted manuscript has been created by University of Washington and the UChicago Argonne, LLC, Operator of Argonne National Laboratory (Argonne). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government. The collection of data used in this study was partly supported by the National Institutes of Health under grant number R01 HD069609 and R01 AG040213, and the National Science Foundation under award numbers SES 1157698 and 1623684.

CONFLICT of INTEREST
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

AUTHOR CONTRIBUTIONS
The authors confirm contribution to the paper as follows: study conception and design: Mohammad Mehdi Oshanreh, Don MacKenzie; analysis and interpretation of results: Mohammad Mehdi Oshanreh; draft

manuscript preparation: Mohammad Mehdi Oshanreh, Don MacKezie, Nazmul Arefin Khan. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Bunch, D. S., and B. Chen. Automobile Demand and Type Choice. In *Handbook of Transport Modelling* (D. A. Hensher and K. J. Button, eds.), Emerald Group Publishing Limited, pp. 541–557.
2. Eluru, N., A. Pinjari, R. Pendyala, and C. Bhat. An Econometric Multi-Dimensional Choice Model of Activity-Travel Behavior. *Transportation Letters*, Vol. 2, No. 4, 2010, pp. 217–230. https://doi.org/10.3328/TL.2010.02.04.217-230.
3. Paleti, R., C. R. Bhat, and R. M. Pendyala. Integrated Model of Residential Location, Work Location, Vehicle Ownership, and Commute Tour Characteristics. *Transportation Research Record*, Vol. 2382, No. 1, 2013, pp. 162–172. https://doi.org/10.3141/2382-18.
4. Bhat, C. R., and J. Y. Guo. A Comprehensive Analysis of Built Environment Characteristics on Household Residential Choice and Auto Ownership Levels. *Transportation Research Part B: Methodological*, Vol. 41, No. 5, 2007, pp. 506–526. https://doi.org/10.1016/j.trb.2005.12.005.
5. JONG, G. D., J. FOX, A. DALY, M. PIETERS, and R. SMIT. Comparison of Car Ownership Models. *Transport Reviews*, Vol. 24, No. 4, 2004, pp. 379–408. https://doi.org/10.1080/0144164032000138733.
6. Basu, R. *Reinterpreting Vehicle Ownership in the Era of Shared and Smart Mobility*. PhD Thesis. Massachusetts Institute of Technology, 2019.
7. Romilly, P., H. Song, and X. Liu. Modelling and Forecasting Car Ownership in Britain: A Cointegration and General to Specific Approach. *Journal of Transport Economics and Policy*, Vol. 32, No. 2, 1998, pp. 165–185.
8. Dargay, J., and D. Gately. Income's Effect on Car and Vehicle Ownership, Worldwide: 1960–2015. *Transportation Research Part A: Policy and Practice*, Vol. 33, No. 2, 1999, pp. 101–138. https://doi.org/10.1016/S0965-8564(98)00026-3.
9. Kitamura, R., and D. S. Bunch. Heterogeneity and State Dependence in Household Car Ownership: A Panel Analysis Using Ordered-Response Probit Models with Error Components. Presented at the Eleventh International Symposium on Transportation and Traffic Theory, Yokohama, Japan, 1990.
10. Whelan, G. Modelling Car Ownership in Great Britain. *Transportation Research Part A: Policy and Practice*, Vol. 41, No. 3, 2007, pp. 205–219. https://doi.org/10.1016/j.tra.2006.09.013.
11. De Jong, G. C. An Indirect Utility Model of Car Ownership and Private Car Use. *European Economic Review*, Vol. 34, No. 5, 1990, pp. 971–985. https://doi.org/10.1016/0014-2921(90)90018-T.
12. Linciano, N. Household Car Usage in the UK. *International Journal of Transport Economics/Rivista internazionale di economia dei trasporti*, 1997, pp. 435–455.
13. Brownstone, D., D. S. Bunch, and K. Train. Joint Mixed Logit Models of Stated and Revealed Preferences for Alternative-Fuel Vehicles. *Transportation Research Part B: Methodological*, Vol. 34, No. 5, 2000, pp. 315–338. https://doi.org/10.1016/S0191-2615(99)00031-4.
14. Choo, S., and P. L. Mokhtarian. What Type of Vehicle Do People Drive? The Role of Attitude and Lifestyle in Influencing Vehicle Type Choice. *Transportation Research Part A: Policy and Practice*, Vol. 38, No. 3, 2004, pp. 201–222. https://doi.org/10.1016/j.tra.2003.10.005.
15. Selvin, S. *Survival Analysis for Epidemiologic and Medical Research*. Cambridge University Press, 2008.
16. Stevens, N.-A., M. Lydon, A. H. Marshall, and S. Taylor. Identification of Bridge Key Performance Indicators Using Survival Analysis for Future Network-Wide Structural Health Monitoring. *Sensors*, Vol. 20, No. 23, 2020, p. 6894.
17. Evrensel, A. Y. Banking Crisis and Financial Structure: A Survival-Time Analysis. *International Review of Economics & Finance*, Vol. 17, No. 4, 2008, pp. 589–602.
18. Holman, D. J., and R. E. Jones. Longitudinal Analysis of Deciduous Tooth Emergence: II. Parametric Survival Analysis in Bangladeshi, Guatemalan, Japanese, and Javanese Children. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, Vol. 105, No. 2, 1998, pp. 209–230.

19. Shankar, V., F. Mannering, and W. Barfield. Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies. *Accident Analysis & Prevention*, Vol. 27, No. 3, 1995, pp. 371–389. https://doi.org/10.1016/0001-4575(94)00078-Z.
20. Abdel-Aty, M. A., and A. E. Radwan. Modeling Traffic Accident Occurrence and Involvement. *Accident Analysis & Prevention*, Vol. 32, No. 5, 2000, pp. 633–642. https://doi.org/10.1016/S0001-4575(99)00094-9.
21. Juan, Z., J. Xianyu, and L. Gao. Understanding Commuters' Daily Travel Time: Application of a Hazard-Based Duration Model. Berlin, Heidelberg, 2011.
22. Bhat, C. R., and J. L. Steed. A Continuous-Time Model of Departure Time Choice for Urban Shopping Trips. *Transportation Research Part B: Methodological*, Vol. 36, No. 3, 2002, pp. 207–224. https://doi.org/10.1016/S0191-2615(00)00047-3.
23. Mohammadian, A., and T. H. Rashidi. Modeling Household Vehicle Transaction Behavior: Competing Risk Duration Approach. *Transportation research record*, Vol. 2014, No. 1, 2007, pp. 9–16.
24. Khan, N. A., and M. A. Habib. Development of Households' Vehicle Choice Models for Travel Accompanying Arrangements. *Transportation Engineering*, Vol. 4, 2021, p. 100071. https://doi.org/10.1016/j.treng.2021.100071.
25. Ghasri, M., T. H. Rashidi, and M. Saberi. Comparing Survival Analysis and Discrete Choice Specifications Simulating Dynamics of Vehicle Ownership. *Transportation Research Record*, Vol. 2672, No. 49, 2018, pp. 34–45.
26. Mannering, F., and C. Winston. A Dynamic Empirical Analysis of Household Vehicle Ownership and Utilization. *The RAND Journal of Economics*, 1985, pp. 215–236.
27. Ramjerdi, F., L. Rand, and I.-A. Sætermo. Models for Car Ownership, Transactions and Vehicle Type. Presented at the URBAN TRANSPORT SYSTEMS. PROCEEDINGS FROM THE 2ND KFB RESEARCH CONFERENCE IN LUND, SWEDEN, 7-8 JUNE, 1999 (BULLETIN 187), 2000.
28. Yamamoto, T., R. Kitamura, and S. Kimura. Competing-Risks-Duration Model of Household Vehicle Transactions with Indicators of Changes in Explanatory Variables. *Transportation Research Record*, Vol. 1676, No. 1, 1999, pp. 116–123.
29. Gilbert, C. C. S. A Duration Model of Automobile Ownership. *Transportation Research Part B: Methodological*, Vol. 26, No. 2, 1992, pp. 97–114. https://doi.org/10.1016/0191-2615(92)90001-D.
30. Mohammadian, A., and E. J. Miller. Dynamic Modeling of Household Automobile Transactions. *Transportation Research Record*, Vol. 1831, No. 1, 2003, pp. 98–105. https://doi.org/10.3141/1831-11.
31. Jensen, A. F., E. Cherchi, and J. de Dios Ortúzar. A Long Panel Survey to Elicit Variation in Preferences and Attitudes in the Choice of Electric Vehicles. *Transportation*, Vol. 41, No. 5, 2014, pp. 973–993. https://doi.org/10.1007/s11116-014-9517-6.
32. Survey Research Center, I. for S. R., University of Michigan. Panel Study of Income Dynamics, Public Use Dataset [Restricted Use Data, If Appropriate]. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, , 2021.
33. Mansournia, M. A., M. Etminan, G. Danaei, J. S. Kaufman, and G. Collins. Handling Time Varying Confounding in Observational Research. *bmj*, Vol. 359, 2017.
34. Cox, D. R. Regression Models and Life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 34, No. 2, 1972, pp. 187–202.
35. Coale, A. J., P. Demeny, and B. Vaughan. *Regional Model Life Tables and Stable Populations: Studies in Population*. Elsevier, 2013.
36. Nelson, W. B. *Applied Life Data Analysis*. John Wiley & Sons, 2005.
37. Kaplan, E. L., and P. Meier. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, Vol. 53, No. 282, 1958, pp. 457–481. https://doi.org/10.2307/2281868.

38. Klein, J. P., and M. L. Moeschberger. *Survival Analysis: Techniques for Censored and Truncated Data*. Springer, 2003.
39. Nelson, W. Hazard Plotting for Incomplete Failure Data. *Journal of Quality Technology*, Vol. 1, No. 1, 1969, pp. 27–52. https://doi.org/10.1080/00224065.1969.11980344.
40. Aalen, O. Nonparametric Inference for a Family of Counting Processes. *The Annals of Statistics*, Vol. 6, No. 4, 1978, pp. 701–726. https://doi.org/10.1214/aos/1176344247.
41. Prillwitz, J., S. Harms, and M. Lanzendorf. Impact of Life-Course Events on Car Ownership. *Transportation Research Record*, Vol. 1985, No. 1, 2006, pp. 71–77. https://doi.org/10.1177/0361198106198500108.
42. Oakil, A. T. M., D. Ettema, T. Arentze, and H. Timmermans. Changing Household Car Ownership Level and Life Cycle Events: An Action in Anticipation or an Action on Occurrence. *Transportation*, Vol. 41, 2014, pp. 889–904.
43. Giuliano, G., and J. Dargay. Car Ownership, Travel and Land Use: A Comparison of the US and Great Britain. *Transportation Research Part A: Policy and Practice*, Vol. 40, No. 2, 2006, pp. 106–124. https://doi.org/10.1016/j.tra.2005.03.002.
44. Ermagun, A., T. Hossein Rashidi, and A. Samimi. A Joint Model for Mode Choice and Escort Decisions of School Trips. *Transportmetrica A: Transport Science*, Vol. 11, No. 3, 2015, pp. 270–289. https://doi.org/10.1080/23249935.2014.968654.
45. Mitra, R. Independent Mobility and Mode Choice for School Transportation: A Review and Framework for Future Research. *Transport Reviews*, Vol. 33, No. 1, 2013, pp. 21–43. https://doi.org/10.1080/01441647.2012.743490.
46. Clark, B., K. Chatterjee, and S. Melia. Changes in Level of Household Car Ownership: The Role of Life Events and Spatial Context. *Transportation*, Vol. 43, 2016, pp. 565–599.
47. Beige, S., and K. W. Axhausen. Interdependencies between Turning Points in Life and Long-Term Mobility Decisions. *Transportation*, Vol. 39, 2012, pp. 857–872.
48. Bhat, C. R., and F. S. Koppelman. A Conceptual Framework of Individual Activity Program Generation. *Transportation Research Part A: Policy and Practice*, Vol. 27, No. 6, 1993, pp. 433–446.
49. Busch-Geertsema, A., and M. Lanzendorf. From University to Work Life – Jumping behind the Wheel? Explaining Mode Change of Students Making the Transition to Professional Life. *Transportation Research Part A: Policy and Practice*, Vol. 106, No. C, 2017, pp. 181–196.
50. Verplanken, B., and D. Roy. Empowering Interventions to Promote Sustainable Lifestyles: Testing the Habit Discontinuity Hypothesis in a Field Experiment. *Journal of Environmental Psychology*, Vol. 45, 2016, pp. 127–134. https://doi.org/10.1016/j.jenvp.2015.11.008.
51. Li, S., and P. Zhao. Exploring Car Ownership and Car Use in Neighborhoods near Metro Stations in Beijing: Does the Neighborhood Built Environment Matter? *Transportation research part D: transport and environment*, Vol. 56, 2017, pp. 1–17.
52. Ding, C., Y. Wang, T. Tang, S. Mishra, and C. Liu. Joint Analysis of the Spatial Impacts of Built Environment on Car Ownership and Travel Mode Choice. *Transportation research part D: transport and environment*, Vol. 60, 2018, pp. 28–40.
53. Esmaili, A., M. M. Oshanreh, S. Naderian, D. MacKenzie, and C. Chen. Assessing the Spatial Distributions of Public Electric Vehicle Charging Stations with Emphasis on Equity Considerations in King County, Washington. *Sustainable Cities and Society*, Vol. 107, 2024, p. 105409. https://doi.org/10.1016/j.scs.2024.105409.
54. Maleki, M., and Smith-Colin, J. (2023). Estimating benefits of microtransit for social determinants of health: A social return on investment system dynamics model. Systems, 11(11), 538.