The Intergenerational Transmission of Automobile Brand Preferences*

Soren T. Anderson
Michigan State University

Ryan Kellogg
University of Michigan

Ashley Langer
University of Arizona

James M. Sallee†
University of Chicago

February 5, 2014

Abstract

We document a strong correlation in the brand of automobile chosen by parents and their adult children, using data from the Panel Study of Income Dynamics. This correlation could represent transmission of brand preferences across generations, or it could result from correlation in family characteristics that determine brand choice. We present a variety of empirical specifications that lend support to the former interpretation and to a mechanism that relies at least in part on state dependence. We then discuss implications of intergenerational brand preference transmission for automakers' product-line strategies and for the strategic pricing of vehicles to different age groups.

*We thank Evan Herrnstadt, Sarah Johnston, and Katie Lim for excellent research assistance and J.P. Dubé, Matthew Gentzkow, Jesse Shapiro, Gary Solon, Raphael Thomadsen, Clifford Winston, and seminar participants at Arizona, Chicago, Copenhagen Business School, Ford Motor Company, Georgetown, Maryland, Michigan, the NBER, the Quantitative Marketing and Economics Conference, Stanford, UC Berkeley, UC Davis, UC San Diego, Washington University, Wayne State, Western Michigan, and Yale for helpful comments.

†Anderson: sta@msu.edu. Kellogg: kelloggr@umich.edu. Langer: alanger@email.arizona.edu. Sallee: sallee@uchicago.edu.

1 Introduction

Economic models generally treat consumer preferences as exogenous and fixed. However, recent studies on tastes for food (Birch 1999; Logan and Rhode 2010; Atkin Forthcoming), female labor supply (Fernández, Fogli and Olivetti 2004), packaged goods (Bronnenberg, Dubé and Gentzkow 2012), and preferences for redistribution (Luttmer and Singhal 2011) suggest that tastes and preferences may form endogenously through prior behavior, consumption, and experience. Such endogeneity has the potential to transform our view of market behavior and public policy by creating a dynamic link between consumption today and welfare tomorrow. For example, long-lived firms considering their pricing strategy today may think ahead to how market share, and consequent brand loyalty, will affect demand among future consumers.

The focus in economics on models of stable exogenous preferences is at least partially pragmatic. Models that allow for endogenous preference formation present technical and empirical challenges, and it is difficult to identify factors that influence preferences in most settings. One influence that is both plausibly important and potentially manageable is the family, which has proven to be fertile ground in prior work (Fernández et al. 2004; Logan and Rhode 2010; Atkin Forthcoming). Families may transmit preferences across generations, not only through immutable genetic endowments, but also through environment and experience. Thus, research on the intergenerational transmission of preferences may provide a better understanding of endogenous preference formation more generally.

In this paper, we empirically investigate the intergenerational transmission of preferences for automobile brands (e.g., Ford) and explore the implications of our results for the automobile market. Automobiles are particularly interesting because they represent a large, infrequent consumer purchase and because the industry remains crucial to the U.S. economy. Brands are interesting in that they represent a "soft" attribute for which we might expect preferences to be especially malleable. That is, it is difficult to see why a consumer would inherently prefer Ford over GM for purely exogenous reasons, such as genetics. Instead, it seems natural to expect such preferences to arise from experience and context.

Our investigation makes use of the Panel Study of Income Dynamics (PSID), which is unique in that it follows multiple households within the same family over time. In particular, it surveys adult children who grew up in a PSID household but have since left and formed their own households. In several recent waves, the PSID has included questions about automobile ownership. Using these data, we find strong correlations in automobile choices across generations within a family. Specifically, a child whose parent has recently purchased a given brand is 39% more likely to choose that same brand (a 5.6 percentage-point increase on a base of 14.3%) than a demographically similar child whose parent did not choose that brand. To the best of our knowledge, we are the first to document this correlation.

We define brand preference somewhat broadly as a situation in which a consumer prefers one automobile brand over another, holding constant major vehicle attributes (such as size and performance) and consumer characteristics (including demographics and geography).[1] This definition allows intergenerational brand preference transmission to be driven by either *intergenerational state dependence* or *direct brand preference inheritance*. In intergenerational state dependence, the brand *choices* of parents influence their children's preferences (and therefore choices) because children have contact with their parents' vehicles and develop tastes for minor design details or nostalgic childhood associations with a brand, or because parental ownership generates information about performance and reliability that is conveyed to children. In direct brand preference inheritance, the brand *preferences* of parents influence their children's preferences (and therefore choices) independently of parental brand choices. For instance, parents might tell their children about their long-term affinity for a brand, or parents might learn about a brand from friends or advertisements and then convey this information to their children prior to purchasing the brand themselves. Since either mechanism can independently create cross-household correlation in brand choices, separate identification of these two mechanisms in our data is difficult. We do find, however, that the correlation between the brand choices of parents and the subsequent choices of their children is stronger for vehicles that were purchased while the children still lived at home and were therefore more exposed to the vehicle. This evidence suggests that intergenerational state dependence is at least one of the mechanisms driving the correlations we observe.

Intergenerational correlation in brand choices could also arise from familial correlations of demographic or geographic factors that determine brand choice. Demographics predict brand choice because consumers with different demographics will prefer different bundles of attributes, and each brand offers a different set of attribute bundles in its products. For example, wealthier households in wealthy families maywill be more likely than households in poorer families to purchase high-end European brands, while larger households will be more likely to choose brands that offer minivans. Similarly, geographic factors, such as terrain andor the local availability ofproximity to dealers and repair shops, maywill also favor certain brands over others. Thus we will need to estimate intergenerational brand loyalty under the identifying assumption that parents'To identify intergenerational brand preference transmission, we require that the unobservable demographics characteristics (apart from brand preference itself) that influence their brand choices of parents are be uncorrelated with children's the unobservable demographics characteristics that influence their brand choices

[1] The distinction between attributes and brands can be blurry, as we discuss below.

of children. This identification problem is challenging because, Iin general, we expect parents and children to share many characteristics, such as wealth, political beliefs, and geographic location, that plausibly influence brand choice. (other than through intergenerational brand loyalty).

We attempt to insulate our estimates from these more mundane brand choice correlations that are driven by these similarities in several four ways. First, we demonstrate that controlling for the rich set of demographic factors available in the PSID has little effect on our estimated brand choice correlations. Second, we show that correlations remain strong when fine geographic controls, which should capture supply factors and consumer sorting, are introduced non-parametrically. FinallyThird, we repeat our analyses for a pair of firms—(GM and Ford) —that produce very similar models with common attributes. offer sets of products such that, for almost any model sold by one firm, a vehicle with a very similar bundle of attributes is available from the other firm. Both brands are also traditional U.S. manufacturers with unionized workforces and similar dealership networks. It is therefore difficult to find demographic factors that would explain households choices between these two brands (indeed, we will show that the observed characteristics in the PSID have essentially no power to explain households choices between Ford and GM), so that the brand choice correlations we observe in this subsample can be more credibly attributed to brand preference transmission. Demographics predict brand choice because consumers with different demographics will prefer different bundles of attributes, and each brand offers a different set of attribute bundles in its products. If two brands offered identical sets of cars, however, then demographics would not predict choice between them—only brand preference would. This greatly mitigates concern about omitted variable bias; in the extreme, if vehicles were so similar that only random events could drive parent's brand choice, then there would be no omitted variable bias. GM and Ford come close to this ideal—they offer sets of products such that, for almost any model sold by one of them, a vehicle with a very similar bundle of attributes is available from the other. Additionally, the two brands have similar non-vehicle attributes that might matter to consumers with different demographics: they are both traditional U.S. manufacturers with unionized workforces and similar dealership networks. We find robust intergenerational brand correlations when examining only the choice between Ford and GM, even while controlling for a rich set offine geographic controls that capture supply side factors. We find similar results for Toyota and Honda. We interpret this as our strongest evidence of true brand preference transmission because it greatlythese comparisons limits the threat of omitted variable bias in our regressions. Fourth, and finally, we show that young adults are particularly influenced by vehicles that were purchased by their parents while they were still living at home. This systematic variation in brand choice correlations is consistent with brand preference transmission—and intergenerational state dependence in particular—rather than mere correlation of characteristics.

Intergenerational brand preference transmission has several implications for automakers' strategies. Such transmission enhances the advantage that brand loyalty gives to incumbent firms because young consumers—who would otherwise be free of loyalty to any firm—arrive at the new car market with preferences inherited from their parents. Furthermore, under the state dependence mechanism, automakers gain loyalty among a future generation of customers when they sell cars to parents, which puts downward pressure on the prices of vehicles targeted to parents. Finally, it is widely believed in the industry that a brand's entry-level vehicles increase firm profits in part by "leading" young consumers to purchase more expensive models from the same firm later in life. In the presence of intergenerational transmission, targeting expensive models to older consumers will also increase profits in the entry-level market by endowing young consumers with brand loyalty, generating a closed loop of benefits to vertical product differentiation. We discuss these results further in the body of the paper, and we flesh out the potential implications for pricing behavior more fully in the appendix, where we derive and calibrate a simple model.

Our analysis relates to several existing literatures. First, as discussed above, our research relates to the nascent literature on endogenous preference formation. Much of that literature centers on preferences for food, which in contrast to automobiles is characterized by small and frequent repeat purchases.

Second, previous work has studied brand loyalty in the automobile market (Mannering and Winston 1985, 1991; Train and Winston 2007). While these papers document within-household brand loyalty, the automobile literature has, to the best of our knowledge, not previously documented intergenerational correlation in brand choice.

Third, our paper relates to the broader literature that studies the magnitude and implications of within-household brand loyalty (sometimes referred to as switching costs) and brand preference persistence. A number of papers, such as Klemperer (1987), Dubé, Hitsch and Rossi (2009), Doganoglu (2010), and Somaini and Einav (Forthcoming), focus on the implications of brand loyalty for equilibrium prices, while others focus on empirically documenting the strength of brand loyalty and brand preferences, typically examining markets for consumer packaged goods (Bronnenberg, Dhar and Dubé 2009; Dubé, Hitsch and Rossi 2010; Bronnenberg et al. 2012). Relative to consumer packaged goods, automobiles are much larger expenses, they are purchased less frequently, and the product offerings are more heterogeneous. Brand loyalty in the automobile sector typically involves individuals purchasing quite different products that share a brand label, whereas the literature on packaged goods is better characterized as repeat purchases of the same item. For small purchases, brand loyalty may be understood as a heuristic to aid in quick decision-making, which is likely quite different from the role that brands play in purchasing an automobile.

Finally, our work has parallels in the extensive peer-effects and social interactions literatures (see Manski 1993, 2000). Whereas much of this literature studies how individuals are influenced by the aggregate behavior and characteristics of a reference group, we focus on how parents and children are influenced by the choices and preferences of a small number of individual family members. Of course, peer effects in automobile purchasing likely extend beyond the family to friends, neighbors, and co-workers. We view parent-to-child preference transmission as a particularly important case of this broader set of peer effects because: (a) parent-child pairs are often more easily identified in data than many other relationships; (b) parent-child relationships have been shown (often using PSID data) to be particularly important for many other economic variables such as income and education (Solon 1992, 1999; Black and Devereux 2011); and (c) our data indicate that parent-to-child transmission is more powerful than that for other within-family links, such as sibling-to-sibling.

The balance of the paper is structured as follows. In section 2 we present a framework for interpreting cross-household correlations in brand choices. We then describe our data in section 3, and we report our empirical results regarding correlations in brand choice across generations in section 4. Section 5 discusses the implications of intergenerational brand preference transmission for automobile markets. Section 6 concludes.

2 Conceptual model of intergenerational vehicle choice

In this section, we present a simple model of household vehicle choice that clarifies possible mechanisms by which choices may be correlated across families and the empirical challenges of separately identifying them. We begin by noting that the distinction between vehicle brands—on which we focus in this paper—and vehicle attributes is blurry. It is tempting to define a brand as something that is independent of all vehicle attributes, as if, for instance, Ford and GM vehicles were identical apart from the logo stamped on the grill. In practice, vehicles of different brands will differ in "minor" features, including trim style, dashboard layouts, and perceived reliability, even for cars that share identical measurable characteristics such as size, power, and cargo space. We define a brand in a way that encompasses these "minor" characteristics so that a brand preference might be derived from, for example, a preference to have the dashboard controls laid out in a particular way. In contrast, when we speak of preferences for attributes, we refer specifically to major vehicle characteristics, such as class, horsepower, size, and fuel economy. We believe that making this distinction between brands and attributes, thusly defined, is useful because the transmission of preferences for brands has a different set of implications than does the transmission of preferences for attributes. The former is primarily relevant for automakers' pricing, marketing, and product line strategies, while the latter is additionally relevant for public policies aimed at addressing the externalities of vehicle use.

Consider a household i in family f that purchases vehicle j at time t. Let the utility that household i derives from this purchase be denoted by:

$$U_{ifjt} = f(D_{ift}, X_j; \beta) + \theta_{ifjt}, \tag{1}$$

where D_{ift} denotes a vector of observed and unobserved demographic and location-specific characteristics of household i, such as income, education, climate, and terrain. These characteristics interact with X_j, which denotes the attributes (including brand) of vehicle j, through the function $f(\cdot)$ and parameter vector β. This interaction allows observable and unobservable characteristics of households and their locations to influence vehicle choice in a variety of ways. For example, rural households may tend to choose pickup trucks, wealthy households may tend to purchase large SUVs, pro-union households may tend to purchase U.S. brands, and households living close to a Ford dealership may tend to purchase Fords. Finally, θ_{ifjt} denotes a preference for vehicle j that is unrelated to demographic or location-specific factors. We focus on influences from other family members as determinants of θ_{ifjt}, but other factors may exist, such as exposure to advertisements, prior driving experiences, idiosyncratic tastes (e.g., for a particular color or trim), or vehicle market conditions at the time of purchase.

Intergenerational brand preference transmission is expressed in our model as a correlation in θ_{ifjt} across households within families, which leads to correlation in vehicle brand choices. Cross-household correlation in θ_{ifjt} could stem from intergenerational state dependence, whereby parental *choices* influence child preferences. For example, if a child's parents purchased a string of GM vehicles, then that child may have nostalgic feelings for GM, a taste for the unique features of GM's design (e.g., the layout of the instrument panel or the feel of the seats), superior information about GM's performance and reliability, or simply a "comfort level" with the brand. Alternatively, correlation in θ_{ifjt} could arise from *direct inheritance* of preferences from parent to child in a way that is not mediated by brand choice itself. For instance, parents might have a belief that Fords have a better-looking trim than do GMs and instill this belief in their children during their childhood. Direct preference inheritance can also occur later in life. For example, parents might be exposed to a positive review of Ford and tell their children about what they read. This mechanism may ultimately lead to a correlation of parents' and children's choices, but it does not operate through state dependence or experience.

Correlations in households' vehicle choices may also arise through cross-household correlations in D_{ift}. It is natural to expect such correlations to exist; Solon (1992), for example, documents strong intergenerational correlation of income. If households with high incomes are more likely to purchase SUVs and European luxury brands, then this correlation in income across generations will lead to correlations in vehicle choices across generations. Thus, a fundamental empirical challenge of our work is to identify vehicle choice correlations that arise from preference transmission (due to either intergenerational state dependence or direct preference inheritance) separately from those that arise from similarities in demographic and geographic characteristics. This identification is important because it is only the former set of channels that is relevant for the strategic implications we consider.

Identification of the "true" transmission of vehicle preferences from parents to children is most clear when there is an exogenous shockexogenous variation that causes the parents to purchase a new vehicle brand, conditional on the parents' and children's demographics. For example, the parents purchase of a Ford may have been driven by a nationwide promotional campaign that coincided with the month in which they wanted to buy a new car , or perhaps It could be that the dealer had a discount when the parents went to test-drive the car or that the weather was particularly nice when the parents test-drove the Ford but stormy when they test-drove the GM.[2] Ideally, we would identify this the "true" transmission of vehicle preferences from parents to children using an instrumental variable that for factors such as these that shifts the choices or preferences of parents but not their children. However, all of our attempts in this direction have been substantially underpowered given our modest sample size.[3] We therefore employ several alternative approaches. First, we leverage the wealth of demographic and location information within the PSID dataset to control directly for potential confounding factors. Despite being able to use a rich set of covariates (including census tract fixed effects), one might nonetheless be concerned that influential unobserved factors remain. For instance, if a family is pro-union, all of that family's households might have a preference for U.S. brands. Thus, we also investigate a subset of the data for which correlated unobserved factors are unlikely to be important: choices between Ford and GM.

[2] These two examples are consistent with the intergenerational state dependence mechanism. Examples consistent with direct brand preference inheritance include the parents reading a positive review of Ford (and then telling their child about the review) or being exposed to idiosyncratic good news about Ford (e.g., Ford vehicles win auto races watched by the parents).

[3] The most obvious instrumental variable is the U.S. market share of the parents' chosen brand at the time of the parents' purchase. However, in an IV version of the regression corresponding to column 2 of table 2, the first-stage F-statistic for this instrument is only 7.06, and the estimated coefficient on parents' brand is -0.127 with a standard error of 0.344. This result compares to an OLS coefficient of 0.093 and standard error of 0.008, as reported below.

Both of these automakers produce a wide range of models with similar attributes and are "iconic" U.S.-based firms, so that it is difficult to imagine demographic characteristics that would drive the choice between one brand or the other. That is, if demographic factors cause an individual to prefer the particular bundle of attributes offered in one of the models made by Ford, it is nearly certain that there is a counterpart made by GM that offers a very similar set of attributes available in the market. Thus, our identifying assumption—that the omitted factors driving the child's brand choice (apart from brand preference itself) are idiosyncratic and not correlated with the unobserved factors driving the parents' brand choice, conditional on observed demographics and location—is more plausible than in the full sample of all brands. In the absence of a viable instrumental variable strategyWhile a powerful, valid instrument would obviously be idealoptimal for identification, we believe that this subsample approach brings us the closest to the ideal setting in whichwhere random eventsidiosyncratic factors push the parents towards one brand rather than another.the other in a way unrelated to the child's inherent preferences.

An even more challenging empirical problem is the separate identification of the extent to which brand preference transmission is driven by state dependence or direct preference inheritance. This distinction is important when considering implications of brand preference transmission for firms' pricing strategies. If only direct preference inheritance is at play, then lowering prices to boost market share among parents will not affect their children's future demand for a brand, whereas the opposite is true if state dependence is at work. This identification problem is similar to one common in the marketing literature, in which one observes a series of brand choices by a single household and then tries to determine whether that household's choices are state-dependent or whether they simply reflect a serially-correlated preference for a particular brand (see, for example, Dubé et al. (2010)). Ideally, we would solve this problem using an instrumental variable, such as vehicle prices, that affects parents' brand choices but not their preferences. However, this strategy is severely under-powered in our setting.[4] Therefore, we adopt the alternative strategy of studying systematic patterns in brand choice correlations that speak differentially to state dependence versus preference inheritance. In particular, sub-section 4.4 studies whether the observed choice correlations are stronger when households have a relatively high level of exposure to the choices of their parents. This systematic variation would be consistent with state dependence but not direct preference inheritance. Moreover, it serves as an additional opportunity to help rule out the possibility that the observed choice correlations are simply an artifact of correlated

[4] Clearly, if the market share instrument discussed above fails to yield sufficient power, a price instrument will fail as well because price effects operate through market shares. Moreover, to the extent that price changes are driven by demand shocks, this instrument may not cleanly separate state dependence from preference inheritance.

demographics between parents and children.

Finally, we note that when we study correlations between children's and parents' purchases, it may be that the child is influencing the parent rather than the other way around. While some of our specifications will explicitly include child-to-parent (and sibling-to-sibling) transmission in the estimated correlations, for the most part we attempt to focus on parent-to-child preference transmission by studying cases in which the parents' purchase preceded their child's. These cases—particularly when we include lagged child's purchases in the regression in section 4.3—isolate the transmission direction for the state dependence mechanism though not necessarily for the direct preference inheritance mechanism.

3 Data

Our data on vehicle ownership come from the Panel Study of Income Dynamics (PSID). In 1968, the PSID surveyed a nationally representative sample of households, and since then it has asked them a battery of economic and demographic questions every year until 1997 and every two years thereafter. The PSID collects information on everyone who lives in a PSID household, but it also follows members of the original PSID sample households and their children whenever they join or create a new household. As a result, the survey now collects information on many households that are members of the same extended family.

The PSID began collecting information on vehicles in 1999. Respondents report the total number of vehicles that they own or lease and additional detailed information on up to three vehicles, including vehicle make, model, and vintage, as well as the date of purchase, purchase price, and whether the vehicle was a gift. These data are available from surveys conducted in 1999, 2001, 2003, 2005, 2007, 2009, and 2011. To the best of our knowledge, the PSID is unique in providing such information for families in the United States.

Our primary focus is on how parental vehicle brand choices correlate with the choices of their adult children. Accordingly, our baseline sample is limited to adult heads of household (or spouses) who purchase a car in the sample and for whom we can identify a parent who owned a vehicle prior to their child's vehicle purchase. We identify 4,338 unique adult children matched to 2,587 unique parents. The difference between the number of parents and children is due to the fact that there are many siblings in our sample.[5]

Table 1 shows sample means for both children and parents in this sample. Adult children are on average 36 years old, whereas parents are 59. Adult children have higher household income, one more year of education (13.5 as opposed to 12.5), and larger household sizes,

[5]In our analysis, we cluster standard errors on the original 1968 PSID family in all regressions to allow for correlated errors across relatives.

Table 1: Variable means and sample sizes in PSID

	Adult children	Parents
Age (years)	36.0	59.4
Years of education	13.5	12.5
Annual family income ($)	78,758	61,171
Number of people in household	3.1	2.3
Number of vehicles owned	2.3	2.1
Number of unique individuals	4,388	2,587

	Matched Pairs
Number of unique child vehicle choices matched to parent choice	16,054
Total number of parent to child vehicle matches	17,268

which accords with the likelihood that they have young children that are still living at home. We observe 16,054 unique vehicle purchases by these 4,388 adult children. Excluded from this sample are vehicles that were received as gifts and vehicles that are likely to have been within-family cross-household sales.[6] In cases where parents are separated, but both are present in the PSID and both have a prior vehicle purchase available, we match the child's vehicle choice with data from both parents.[7] There are 1,214 such cases, which gives us 17,268 parent-child vehicle pairs in our main estimation sample.

4 Empirical evidence of intergenerational brand preference transmission

In this section, we develop and estimate a linear probability model (LPM) of the relationship between brand choices of children and the choices of their parents, as well as other covariates. We employ a LPM rather than a structural discrete choice model because it is more forgiving to the extensive geographic and time fixed effects that we use in our estimation, though this comes at the cost of not being able to interpret our coefficient estimates as parameters of a utility function.

To operationalize our brand choice data in the LPM framework, we first categorize all vehicle choices as being one of seven "brands": GM, Ford, Chrysler, Toyota, Honda, Other Asian, and European. Grouping smaller Asian automakers and European manufacturers together ensures that each brand is chosen frequently enough to yield meaningful estimates

[6]Specifically, we drop a child's vehicle purchase if the parent household owned the same make, model, and model-year in the previous survey wave and subsequently no longer owns the vehicle following the child's purchase.

[7]In these cases we weight each vehicle-parent pair by half so that choices that appear twice in our data are weighted equally to those that appear once.

in a linear probability framework (these brand definitions imply that all choice probabilities lie in the 4%–33% range in the raw data).[8]

We build a linear probability model with multiple choice possibilities by stacking a set of binary linear probability models for each of our 7 brands. To motivate our approach, we first consider a linear probability model for a single brand. Our hypothesis is that parental ownership of a given brand will make a child more likely to choose that brand. For example, to test this hypothesis for Ford we could run a linear probability model in which the dependent variable is coded as one if the child was observed to choose a Ford. In addition to controls for the child's demographics and parents' demographics, our regressor of interest would be a dummy variable for the parents' brand choice, which would be coded as 1 if the parents' most recent vehicle purchased prior to the child's purchase was a Ford.[9] The one-brand estimation equation would be:

$$Ford_{ift} = \gamma \cdot Ford_{pft} + \mathbf{X_{ift}}'\beta + \mathbf{X_{pft}}'\delta + \alpha_t + \epsilon_{ift}, \tag{2}$$

where $Ford_{ift}$ is coded as 1 if child i of family f at time t chose a Ford, $Ford_{pft}$ is coded as 1 if parent p's most recent choice before t was a Ford, α_t is a period-specific constant to capture Ford's average market share in each period, $\mathbf{X_{ift}}$ are the child's observable demographic characteristics, and $\mathbf{X_{pft}}$ are those of the parents. In this regression of "Ford against the field", all observations in our data are included.

Instead of running 7 separate one-brand linear probability models, we stack them and run pooled OLS. For each of our seven brands, we create an observation for each observed choice situation (each car purchase) in which the dependent variable is coded as 1 if the brand is chosen and zero otherwise. We do this for all seven brands and stack the data, which generates a final data set that has seven times the number of observations as our original data set that included one observation per choice.[10]

[8] In order to test whether the correlation across generations is coming from a correlated preference for brand (e.g., Ford) or sub-brand (e.g., Ford, Lincoln, or Mercury), we have run subsets of our regressions with 41 sub-brands instead of the 7 brands, interacting our control variables with all 41 sub-brands. In general, we find that both the overall brand and the sub-brand of the parent have a statistically significant correlation with the sub-brand chosen by the adult child.

[9] Year of purchase is recorded for all but a handful of very old cars, which we dropped. Month of purchase is missing for 37% of all child and 31% of all parent purchases, however, including all purchases in the 2009 and 2011 waves. In establishing whether a parent vehicle was purchased before a child vehicle, we treat these observations conservatively by coding parent purchase months as December and child purchase months as January.

[10] Our procedure can also be described as expanding each observed choice as follows. For each vehicle purchase by every individual in our data, we expand the original data sample to include seven lines of data. The first is for the brand that was chosen by the individual, and this line has the dependent variable coded as one. The other six are observations with a zero dependent variable, one for each of the six brands not chosen.

We interact all of our regressors with brand dummies, thereby allowing observed demographic and geographic factors to affect the choice probability of each brand differently. Thus, all covariates are denoted with a j subscript or interacted with a coefficient vector of length j. The one restriction that we impose is that the effect of the parents' past brand choice is common across brands.[11] Thus, instead of seven dummy variables for parental brands, there is only one dummy variable coded as 1 when the parental choice matches the brand represented in the corresponding row of data for the child. For each child's purchase, exactly 1 of 7 observations will have the parental dummy variable coded as 1.

This leads us to the following estimation equation:

$$b_{ifjt} = \gamma \cdot 1(b_{pfjt} = b_{ifjt}) + \mathbf{X_{ift}}'\beta_j + \mathbf{X_{pft}}'\delta_j + \alpha_{jt} + \epsilon_{ifjt}, \tag{3}$$

where the dependent variable, b_{ifjt}, is a dummy coded as 1 if child i of family f chose brand j in choice t. The independent variable of primary interest is a dummy variable that indicates whether the parents' most recent prior purchase is of that same brand: $1(b_{pfjt} = b_{ifjt})$.[12]

Our hypothesis is that the γ coefficient will be positive; that is, children are more likely to purchase a given brand if their parents have purchased that brand in the recent past. We control for both child characteristics X_{ift} and parent's characteristics X_{pft}, which enter with brand-specific coefficient vectors, β_j and δ_j, that we estimate by interacting child and parent characteristics with brand dummies. Finally, we allow for brand-by-month of purchase fixed effects, α_{jt} to capture overall market shares, leaving ϵ_{ifjt} as the error term.

This setup expands each observed brand choice into seven observations. To avoid this expansion of the data set unduly shrinking our standard error estimates, we cluster all standard error calculations at the level of the 1968 PSID family. This clustering accounts for the mechanical correlation in the residuals between the seven observations that represent a single choice, the correlation in each individual's choices across choice situations, and the correlation across siblings or cousins. We also weight each observation using PSID-provided

[11] By using the linear probability model, we do not impose a restriction that predicted values must be between zero and one, nor do we require that the sum of the predicted values across the seven brands must equal one for each choice situation. We have checked our predicted values for our baseline specifications, and we find that the vast majority of predicted values are between zero and one, and those that deviate are very small negative numbers. Similarly, the sum of the predicted values across the seven brands for each choice situation are tightly distributed around 1.

[12] We have experimented with a variety of ways of characterizing the parents' choice given that parents may own multiple vehicles. In section 4.3 below, we include additional lags of the parents' choices. In section 4.4, we expand the dataset so that each child's purchase is matched to all of the parents' prior purchases, not just the most recent purchase. We have also experimented with an independent variable measuring the share of parents' vehicles from a specific brand, as well as with matching child and parent vehicles one-to-one based on the order in which the vehicles are listed in the survey. In all cases, our qualitative results are quite similar.

sampling weights so that the original PSID households on which our sample is based can be interpreted as representing the U.S. population at the time of the original survey.

We do not allow for an outside good, which would be interpreted as the option to not purchase a vehicle at all. Inclusion of an outside good is standard in discrete choice modeling, but here we are interested in knowing whether or not a child, conditional on purchasing a vehicle, decides to buy a brand that is the same as the one owned by members of his or her family. Inclusion of an outside good would conflate correlations in choice that determine whether or not individuals purchase vehicles with correlations in the brand chosen when purchasing a vehicle, which are distinct phenomena.

4.1 Baseline results

We begin by showing simple correlations in order to demonstrate the strength of the intra-family relationship and then demonstrate how the correlation is affected by various controls. We focus here and in section 4.2 on separating intergenerational brand preference transmission from choice correlation driven by demographic or geographic factors. Section 4.3 then studies whether the choice correlations are caused by short-run or long-run mechanisms, and in section 4.4 we attempt to distinguish between the intergenerational state dependence and direct brand preference inheritance mechanisms.

Table 2 presents coefficient estimates of γ from equation 3, which regresses the brand chosen by the child on a dummy for whether or not the parents' most recent purchase is the same brand. The estimate in column 1, which includes only month-of-purchase by brand fixed effects that control for the overall share of each brand during each period, indicates that a child is 9.6 percentage points more likely to choose the brand that her parents chose. There are seven brands in our choice set, so the probability that the average brand is selected is 0.143. Thus, our estimate implies that a child whose parents chose a particular brand is 67% more likely to choose that brand than another child whose parents chose differently.

This is a remarkably strong relationship, but it may reflect not only the intergenerational transmission of brand preference that we are interested in, but also familial correlations in demographic and location-specific factors that cause related households to demand similar attributes in vehicles, in turn causing a correlation in brand choice.

As a first step toward addressing this issue, we introduce progressively richer controls in columns 2 through 7 of table 2 and examine how the coefficient estimates change. In column 2 we add demographic controls (including family income, age, sex, education, number of kids in the household, and household size) for the child's household and the parents' household.[13]

[13]Each of these characteristics is interacted with a dummy for each brand, which is a flexible analog to

Table 2: Correlations between child's brand choice and parents' brand choice

	Dependent Variable: Child's Brand						
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Parents' brand = child's brand	0.096	0.088	0.077	0.075	0.056	0.050	0.038
	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.012)	(0.013)
Month of purchase fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Child's demographics	No	Yes	Yes	Yes	Yes	Yes	Yes
Parents' demographics	No	Yes	Yes	Yes	Yes	Yes	Yes
Child's state fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
Parents' state fixed effects	No	No	No	Yes	Yes	Yes	Yes
Child's county fixed effects	No	No	No	No	Yes	Yes	Yes
Child's census tract fixed effects	No	No	No	No	No	No	Yes
Number of choices	17,268	17,268	17,268	17,268	17,268	6,937	6,937
R^2	0.085	0.096	0.113	0.119	0.216	0.280	0.416

Standard errors clustered by 1968 PSID family are in parentheses. Each column is a linear probability model where each individual-year-vehicle choice enters the data 7 times, once for each brand (GM, Ford, Chrysler, Toyota, Honda, Other Asian, and European). Child's and parents' demographics include age, education, income, gender, number of children in household, and family size. All control variables are interacted with 7 dummies, one for each brand. Columns 6 and 7 limit the sample to households living in census tracts that contain more than one PSID family.

The addition of these controls lowers the estimated coefficient from 0.096 to 0.088. While the modest impact of these controls on the estimated coefficient is encouraging, the regression's R^2 also changes only slightly, from 0.085 to 0.096. Following the logic of Altonji, Elder and Taber (2005), the fact that this modest increase in explained variation yields a measurable drop in the estimated coefficient raises the possibility that, if unobserved demographic factors that affect children's brand choices are as correlated with parents' choices as are the observed demographic factors, then these unobservables may explain much of the observed brand choice correlation. We further address this issue in section 4.2 where we focus on two brands—Ford and GM—that have very similar attributes so that choices between them are therefore unlikely to be affected by correlated demographic factors.

The remaining columns of table 2 address common geographic factors that might lead to choice correlation. Column 3 adds state-by-brand fixed effects, which control for differences in market shares and location-specific factors that vary by state. Column 4 adds analogous fixed effects for parents. These fixed effects cause the estimated coefficient to fall from 0.088 to 0.075.

Geographic factors, such as dealer location, local prices, weather, and terrain, may vary

the traditional approach in the automobile demand literature of interacting vehicle attributes with buyer characteristics. We additionally experimented with a large assortment of additional financial controls from the PSID, such as amounts spent on vacations, eating out, health insurance, clothes, and a variety of other expenditures and found that these did not affect the estimates substantially.

significantly within some states. In column 5 we add county-by-brand fixed effects for the child's county of residence, which is intuitively a small enough geographic area to control for most omitted factors that we have in mind. County fixed effects lower the point estimate to 0.056, which is approximately 58% of the magnitude of the raw correlation and still highly significant, both statistically and economically. This estimate implies that parental ownership boosts the conditional probability that a child buys a given brand by 39%.

We interpret the difference across columns in the estimated effects as evidence that some location-specific factors are both important in determining brand choice and correlated across family members (who tend to live in similar places). Weather, terrain, urbanization, and culture are important determinants of the demand for attributes, which are different on average across brands. Even conditional on demand, there may be a different availability of brands across geographic areas due to the location of dealerships. In our view, most of these differences should be captured by county-level fixed effects, so column 5 is our preferred specification.

Nonetheless, we can push further and include fixed effects for each census tract (a unit of approximately 2,500 to 8,000 people) in the sample to address very fine-scale local factors (local repair shops, for example). We are able to include these fixed effects because the original PSID sample design drew stratified samples from particular geographic areas. The legacy of that original sample design is that PSID households are still more geographically clustered than would be the case for a random sample of households. That said, a large number of census tracts in our sample hold only one PSID family, so to use census tract fixed effects we must first restrict the sample to tracts in which we observe multiple families.[14] Column 6 re-estimates the model with county level fixed effects on this subsample, and column 7 estimates a model with census tract fixed effects. Even when including these effects, our coefficient of interest is economically and statistically significant. The census tract fixed effects are powerful, increasing the R^2 from 0.280 to 0.416 and further lowering the point estimate between columns 6 and 7 (though the coefficients are still comfortably within each other's confidence intervals).[15]

Our baseline regressions focus on parental choices determining child choices, which we believe to be the strongest intrafamily channel of brand preference transmission. We can, however, configure our data to examine the relationship that prior purchases by any family

[14]If we do not restrict the sample in this way, the main coefficient is then primarily identified off of within-family variation in brand choice over time rather than cross-family variation. This identifying variation largely excludes long-term effects from the estimate so that it cannot be compared to the estimates in columns 1 through 5 (see section 4.3 for a fuller discussion of long-run versus short-run effects).

[15]When we restrict our attention to similar brands in section 4.2, this decrease in the point estimate is substantially mitigated.

member have to subsequent choices by their relatives. To do so, we take every vehicle choice observed in the data and match it to the most recent purchase made by every other related household in the dataset (including parents, children, siblings, cousins, etc.). We then include all of these bilateral relationships in one regression, down-weighting vehicles that are matched to multiple family members' vehicles so that they have equal influence on the estimate as those that have only one match. This alternative construction expands our sample size considerably and delivers more precise, but modestly smaller effects. For example, the all-family matched analog of column 1 from table 2 produces a coefficient (standard error) estimate of 0.068 (0.005), and the county fixed effects analog to column 5 produces an estimate of 0.030 (0.004). These results are consistent with our intuition that parent-to-child influences are particularly strong, but it also suggests that broader family network effects have influence.

4.2 Estimates limited to similar brands

The principal concern with our baseline regressions is that demographic or location-specific characteristics of children and parents will be correlated and that these characteristics drive demand for vehicle attributes that are correlated with brand. While we believe that controlling for county and census tract fixed effects adequately addresses location-specific confounders, the possibility remains that our estimates are contaminated by demographic confounders, even after controlling for observables. For example, individuals who work in construction occupations may be more likely to have children that work in construction, and both the parents and children may thustherefore prefer light trucks to passenger cars. and to prefer light trucks to passenger cars. Because GM's fleet is more heavily tilted toward light trucks than is Honda's, such peopleboth the parents and children will then be more likely to buy a GM, as will their children, even in the absence of any brand preference transmission.

Here, we address this issue by isolating the choice set to two brands that are very similar: Ford and GM. Ford and GM are both full-line, U.S.-based automakers that compete directly in every vehicle segment.[16] Because their vehicle lineups are so similar, conditional on local supply, we expect that random events , idiosyncratic variation willto largely drive the choice between Ford and GM in the absence of brand preference. We therefore anticipate that intrafamily brand choice correlations would be quite weak in the absence of intrafamily brand

[16]Popular perception holds that Ford and GM are similar brands. We have confirmed this empirically using a measure of "distance" between brands that borrows from Langer and Miller (2013), who calculate the distance between pairs of vehicles in attribute space based on vehicle segment, price, number of passengers, wheelbase, fuel economy, and horsepower for GM, Ford, Toyota and Chrysler. Using their metric, we have confirmed that Ford and GM vehicles are on average substantially closer to each other than they are to Toyota or Chrysler.

Table 3: Correlations between child's brand choice and parents' brand choice among those owning a Ford or GM

	Dependent Variable: Child's Brand						
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Parents' brand = child's brand	0.135	0.137	0.115	0.111	0.084	0.070	0.065
	(0.021)	(0.021)	(0.019)	(0.019)	(0.019)	(0.030)	(0.034)
Month of purchase fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Child's demographics	No	Yes	Yes	Yes	Yes	Yes	Yes
Parents' demographics	No	Yes	Yes	Yes	Yes	Yes	Yes
Child's state fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
Parents' state fixed effects	No	No	No	Yes	Yes	Yes	Yes
Child's county fixed effects	No	No	No	No	Yes	Yes	Yes
Child's census tract fixed effects	No	No	No	No	No	No	Yes
Number of choices	9,355	9,355	9,355	9,355	9,355	3,587	3,587
R^2	0.074	0.076	0.100	0.113	0.268	0.307	0.452

Standard errors clustered by 1968 PSID family are in parentheses. Sample is limited to the cases where the child chose Ford or GM. Each column is a linear probability model where each individual-year-vehicle choice enters the data 7 times, once for each brand (GM, Ford, Chrysler, Toyota, Honda, Other Asian, and European). Child's and parents' demographics include age, education, income, gender, number of children in household, and family size. All control variables are interacted with 7 dummies, one for each brand. Columns 6 and 7 limit the sample to households living in census tracts that contain more than one PSID family.

preference transmission when we limit our sample to children who choose either a Ford or a GM. That is, the unobserved demographic variables that are likely to be correlated between parents and children are unlikely to drive the choice between these two brands.

Table 3 repeats the specifications in table 2 for a subset of choices limited to Ford and GM. Specifically, we keep all instances in which a child chose either a Ford or GM, which accounts for about 54% of our original sample.[17] As in the full sample, the results are all positive, statistically significant, and economically large, corroborating our baseline results and casting doubt on the possibility that the correlation in brand choice across households is due entirely to demographic confounders. Notably, the estimated coefficient actually increases slightlyhardly changes going from column 1 to column 2 (it actually increases slightly) as the demographic controls are added, which was not the case in the "all brands" specification from table 2. This result comports with the intuition that demographic factors are unlikely to influence the choice between Ford and GM. Moreover, , following the logic of Altonji et al. (2005), if our "headline" demographic variables are not correlated with the choice between Ford and GM, it seems unlikely that unobserved factors would be correlated

[17]We do not restrict the sample based on whether the parents chose Ford or GM. We do, however, add an additional control variable for whether the parents' choice was one of these two brands. This control helps ensure that the sum of the child's choice probabilities for the two brands is close to one.

either. It is also worthwhile to note that when the county fixed effects are replaced with census tract fixed effects going from column 6 to column 7, the estimated coefficient hardly changes despite a substantial increase in R^2, suggesting that county-level effects are sufficient to control for local factors affecting brand choice.[18]

The magnitudes of the coefficients are somewhat larger in the subsample in table 3 than in the full sample in table 2. Nonetheless, parents' brand choice has a slightly smaller percentage effect on child's choice probabilities in the restricted sample because the baseline choice probabilities are higher. In the full sample, the market share for Ford is 22%, while that for GM is 33%. In the subsample, the corresponding figures are 39% for Ford and 61% for GM. Thus, the coefficient of 0.084 in column 5 of table 3 (the county fixed effect specification in the restricted sample) implies that a child whose parents' most recent prior purchase was the same brand boosts the probability of purchase by 22% for Ford and 14% for GM, whereas the analogous coefficient in table 2 (the full sample) represents a 25% effect for Ford and 17% for GM. This difference is intuitive given that Ford and GM are generally close substitutes.

Toyota and Honda are also similar brands. They both produce a full range of sedans and fuel efficient SUVs, though Honda produces only a limited set of pickup trucks. Table 4 shows results from the same set of specifications for the sample of observations limited to children who purchased either a Honda or a Toyota, excluding all pickup trucks. The estimated effects in this subsample are even larger; the county fixed effects specification in column 5 (our preferred specification) indicates that having a parent who owns a Honda or Toyota increases the probability that a child chooses that brand by 63%.[19] Honda and Toyota have a smaller market share than Ford and GM, which leaves us with a smaller sample size and larger standard errors. Nevertheless, our estimates are statistically significant at any conventional level except when we limit the sample to those who live in a census tract common to another PSID family in our sample (columns 6 and 7), at which point we lose power.

4.3 Long-run versus short-run effects

In this section, we examine the extent to which the brand choice correlations documented above are driven by long-run or short-run brand preference transmission. Both long-run and short-run transmission could be associated with the intergenerational state dependence

[18]As an additional test for local confounders, we have estimated the Ford/GM specification using a subsample in which the child lives in a different state than the parents (including demographics and state fixed effects in the specification, for both child and parent). The point estimate of γ from this subsample is 0.139, which is actually larger than the estimate of 0.111 from column 4 of table 3.

[19]A 31.4 percentage point increase on an average market share of 50% is a 63% increase.

Table 4: Correlations between child's brand choice and parents' brand choice among those owning a Honda or Toyota

	Dependent Variable: Child's Brand						
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Parents' brand = child's brand	0.228	0.233	0.265	0.269	0.314	0.500	0.549
	(0.066)	(0.066)	(0.063)	(0.065)	(0.071)	(0.353)	(0.511)
Month of purchase fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Child's demographics	No	Yes	Yes	Yes	Yes	Yes	Yes
Parents' demographics	No	Yes	Yes	Yes	Yes	Yes	Yes
Child's state fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
Parents' state fixed effects	No	No	No	Yes	Yes	Yes	Yes
Child's county fixed effects	No	No	No	No	Yes	Yes	Yes
Child's census tract fixed effects	No	No	No	No	No	No	Yes
Number of choices	2,327	2,327	2,327	2,327	2,327	352	352
R^2	0.116	0.126	0.160	0.200	0.469	0.826	0.909

Standard errors clustered by 1968 PSID family are in parentheses. Sample is limited to the cases where the child chose Honda or Toyota. Each column is a linear probability model where each individual-year-vehicle choice enters the data 7 times, once for each brand (GM, Ford, Chrysler, Toyota, Honda, Other Asian, and European). Child's and parents' demographics include age, education, income, gender, number of children in household, and family size. All control variables are interacted with 7 dummies, one for each brand. Columns 6 and 7 limit the sample to households living in census tracts that contain more than one PSID family.

mechanism or the direct preference inheritance mechanism. For instance, long-run transmission could derive from early-childhood experience with a brand or from parents' repeated statements to a child that they believe one brand is better or more reliable than another. Short-run transmission could come from information about parents' recent purchases or from parent-child discussions about recent reviews of or advertisements by a particular brand.

We begin by studying how children's brand choices correlate with multiple previous choices of their parents, including both the most recent choices as well as earlier choices. To do so, we augment our baseline regression with indicator variables for whether the brand of each child's purchase matches the brand of the parents' lagged purchases.[20]

Ideally, we would like to use parents' complete life history of vehicle ownership to test how vehicles owned at different points in their children's lives influenced the children's subsequent choices. Unfortunately, we are limited to seven waves of data, which leaves us with relatively few car purchases for most families and a complete history for none. Thus, we are only able estimate and compare how children's brand choices correlate with the most recent versus somewhat less recent choices of their parents. If children's brand choices correlate more

[20] Observations of lagged purchases are not available for all child's purchases. Rather than drop observations that are missing lags, we also include in our regressions interactions between the brand dummies and indicator variables (one for each lag in the regression) that equal one if the lag is missing.

strongly with their parents' most recent choices, then short-run preference transmission is likely important. If less recent parent choices still have predictive power conditional on recent parent choices, then longer-run preference transmission is also likely important.[21]

When we add parents' lagged brand choices to the regression, we also add indicators for the children's own lagged brand choices. We do so for two reasons. First, in the presence of within-household state dependence, the lagged choices of a parent may continue to influence a child's current choice indirectly via the child's earlier brand purchase. Thus, to test whether lagged parent purchases have a direct long-run influence on child choices, we must control for the child's lagged purchases.

Second, these lags help, in part, to identify parent-to-child transmission separately from child-to-parent transmission. In particular, children's choices might have a short-run influence on the preferences of their parents—the state dependence mechanism from child to parent. If so, then this correlation could propagate via within-household state dependence to generate a longer-term correlation between recent child choices and lagged parent choices. Conditioning directly on the lagged choices of children helps protect against this concern.[22] Unfortunately, these lags do not address the possibility that children might influence their parents through the direct preference inheritance mechanism. For instance, a child might tell her parents about a recent blog post arguing that Fords are a great buy, leading both the parents and then the child to buy a Ford. In this case, child and parent choices will be correlated, even controlling for the child's lagged purchases.[23]

Table 5 reports linear probability model regressions that include indicators for parents' and children's lagged purchases. All regressions use our preferred specification with demographic controls and county fixed effects. Columns 1 through 3 use all seven brands and include, progressively, zero, one, and two lags of both the parents' and children's purchases; columns 4 through 6 repeat these specifications for the Ford and GM subsample. The results show little difference between the estimated coefficients on the parents' most recent and lagged choice. In the Ford and GM subsample, the lagged coefficient is actually larger than the most recent coefficient, though the two are not statistically distinct ($p = 0.527$ in

[21]For reference, the average time elapsed between the child's purchase and the parents' lagged and second lagged purchases are 40 months and 65 months, respectively. Thus, even our limited specification is suggestive of effects that take place over several years.

[22]Of course, child-to-parent state dependence could arise from more distant child's purchases than the second lag. We have also estimated specifications that include third and fourth lags of children's purchases and found that doing so has only a minor effect on the estimated coefficients. For instance, adding third and fourth own-lags to specification (3) of table 5 below reduces the coefficient on the indicator for parents' brand = child's brand from 0.040 to only 0.038.

[23]We focus here on short-run preference transmission, since it seems unlikely that young children still living with their parents would be able to influence the longer-run brand preferences of their parents in this way.

Table 5: Long-run and short-run correlations between child's brand choice and parents' brand choice

	Dependent Variable: Child's Brand					
		All brands			Ford and GM only	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Parents' brand = child's brand	0.056	0.042	0.040	0.084	0.059	0.058
	(0.007)	(0.006)	(0.006)	(0.019)	(0.017)	(0.016)
Lagged parents' brand = child's brand		0.044	0.040		0.080	0.075
		(0.007)	(0.007)		(0.020)	(0.020)
2nd lagged parents' brand = child's brand			0.018			0.013
			(0.008)			(0.024)
Lagged child's brand = child's brand		0.131	0.124		0.225	0.214
		(0.007)	(0.007)		(0.020)	(0.019)
2nd lagged child's brand = child's brand			0.068			0.092
			(0.007)			(0.023)
Month of purchase fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Child's demographics	Yes	Yes	Yes	Yes	Yes	Yes
Parents' demographics	Yes	Yes	Yes	Yes	Yes	Yes
Child's state fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Parents' state fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Child's county fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Child's census tract fixed effects	No	No	No	No	No	No
Number of choices	17,268	17,268	17,268	9,355	9,355	9,355
R^2	0.216	0.229	0.232	0.268	0.293	0.297

Standard errors clustered by 1968 PSID family are in parentheses. Each column is a linear probability model where each individual-year-vehicle choice enters once for each brand. Child's and parents' demographics include age, education, income, gender, number of children in household, and family size. All control variables are interacted with dummies for each brand.

column 6). Parents own 2.1 vehicles on average (see table 1), so the similarity of the most recent and first lagged brand choice of parents may reflect the fact that the two most recent purchases on average represent the parents' current fleet. The estimated coefficients on the parents' second lag are, however, substantially smaller than those on the parents' most recent purchase. This difference is statistically significant (p = 0.031) for the all brands sample and marginally significant (p = 0.116) for the Ford/GM subsample. The estimated coefficients on the parents' second lagged choices themselves are positive (and statistically significant in the all brands sample). Overall, these results suggest that both short-run and long-run preference transmission are at work.[24]

[24] We are limited in how many lags we can include by data availability. Another approach is to proxy for the parent's past purchases by including directly the market share of each brand in the parent's geographic region further back in time. We do not, however, have market share data at the subnational level for most years. We can construct market shares for 1990 at the level of the Metropolitan Statistical Area (MSA) using the National Household Transportation Survey. When we include the 1990 market share of the brand in the parents' MSA as an additional control variable, we find that these market shares are positively correlated with child choice, but statistically imprecise.

We next attempt to isolate the short-run mechanism by examining regression models that include a fixed effect for each child household. These fixed effects account for any household-level permanent brand preference so that the only source of identification comes from changes in parents' brand choices over time. For these regressions, we revert to our original specification that includes only the parents' most recent brand choice.

Results from the household fixed effects specification are given in table 6, in which all columns include demographic controls. Column 1 includes all seven brands, while column 2 uses only the Ford and GM subsample to guard against time-varying unobservable demographic factors that may be correlated between child and parent. In both of these specifications, the estimated coefficient of interest is positive but statistically insignificant. To improve the estimates' precision, we expand the sample to include all bilateral relationships that are available in the data (that is, including siblings, uncles, cousins, etc. as discussed at the end of section 4.1 above). When we use this expanded sample, the estimated correlation between children's choices and relatives' choices is positive and statistically significant, as shown in columns 3 and 4 of table 6.[25] While the magnitudes of these coefficients are small relative to those obtained without household fixed effects (the corresponding point estimates are 0.062 for column 3 and 0.102 for column 4, still using the "all relatives" sample), they nonetheless provide evidence that brand preference transmission has a short-run component and is not entirely driven by childhood experiences.

4.4 Tests for intergenerational state dependence

In this section, we explore the extent to which we can distinguish intergenerational state dependence from direct preference inheritance. To do so, we use variation in the exposure of children to their parents' vehicles, under the logic that an increase in exposure should strengthen the state dependence mechanism but not the direct preference inheritance mechanism. Specifically, we study whether children are more strongly influenced by vehicles that their parents owned while they still lived with their parents, under the presumption that children were more likely to have direct exposure to such vehicles.

For this analysis, we expand the sample by matching each child's choice to all observed choices by their parents that precede the child's choice, not just to the parents' most recent purchase. Thus, for each vehicle that a child purchased, there may be multiple observations in the regression sample, one for each parent purchase.[26] For each matched child and parent

[25]We have also run the column 4 specification while including county fixed effects to account for households that move during the sample. The results are robust to these fixed effects: the estimated coefficient and standard error are 0.016 and 0.006, respectively. We have not run the column 3 specification with county fixed effects due to computer memory constraints.

[26]To be clear, we do not add parents' past brand choices as lagged regressors, as we did in section 4.3

Table 6: Correlations between child's brand choice and relative's brand choice, including household fixed effects

	Dependent Variable: Child's Brand			
	Match to parents only		Match to all relatives	
	All brands	Ford and GM	All brands	Ford and GM
VARIABLES	(1)	(2)	(3)	(4)
Relative's brand = child's brand	0.007	0.038	0.006	0.020
	(0.009)	(0.028)	(0.002)	(0.007)
Month of purchase fixed effects	Yes	Yes	Yes	Yes
Child's demographics	Yes	Yes	Yes	Yes
Relative's demographics	Yes	Yes	Yes	Yes
Child fixed effects	Yes	Yes	Yes	Yes
Child's state fixed effects	No	No	No	No
Relative's state fixed effects	No	No	No	No
Child's county fixed effects	No	No	No	No
Child's census tract fixed effects	No	No	No	No
Number of choices	17,268	9,355	126,893	71,775
R^2	0.462	0.591	0.477	0.593

Standard errors clustered by 1968 PSID family are in parentheses. Each column is a linear probability model where each individual-year-vehicle choice enters once for each brand (GM, Ford, Chrysler, Toyota, Honda, Other Asian, and European). Child's and Relative's demographics include age, education, income, gender, number of children in household, and family size. All control variables are interacted with dummies for each brand.

purchase, we then identify whether the parents' purchase was made while the child was still living in the parents' household. There are 16,115 such "child at home" purchases made by parents in the data, out of 58,420 total parent purchases. On average, the children in this subsample are 16.9 years old while living at home when their parents buy a vehicle (standard deviation = 5.5 years) and 26.0 years old after having moved out when they buy their own vehicle (standard deviation = 4.6 years). The average age at which children matched to "at home" cars move out of their parents' home is approximately 22.4 years old.[27]

Our primary variable of interest is an interaction between the indicator for whether the parents' brand choice matches the child's brand choice with an indicator for whether the parents' purchase occurred while the child was still at home. Our idea is that children will

above. Instead, we create a separate observation for each child-parent match. We do so for two reasons. First, the number of available parent purchases varies substantially across child households. Second, this approach eases the interpretation of the interacted regressors discussed below. We re-weight the observations in this expanded sample such that every case of a child's vehicle choice receives equal weight, regardless of how many parent vehicles it was matched to; we then apply the PSID sampling weights. As before, we ultimately expand this sample according to the number of brands in our linear probability model framework.

[27]We do not know the exact date at which children moved out; we only know the survey date at which each child is observed in his/her own household. The average age at this observation is 23.4; we subtract one year under the assumption that move-out dates are uniformly distributed across the two years between surveys.

Table 7: Correlations between child's brand choice and parents' brand choice interacted with whether child was living with parents

	Dependent Variable: Child's Brand					
		All brands			Ford and GM only	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Parents' brand = child's brand	0.0316	0.0364	0.0535	0.0191	-0.0342	-0.0285
	(0.0152)	(0.0228)	(0.0259)	(0.0390)	(0.0727)	(0.0799)
(Parents' brand = child's brand)	0.0144	0.0138	0.0148	0.0608	0.0683	0.0687
x child at home	(0.0118)	(0.0114)	(0.0114)	(0.0299)	(0.0300)	(0.0310)
(Parents' brand = child's brand)	4.52e-04	1.05e-04	-2.44e-03	0.00144	0.00540	0.00453
x child's age at purchase	(4.86e-04)	(1.52e-03)	(2.79e-03)	(0.00130)	(0.00507)	(0.00907))
(Parents' brand = child's brand)		5.69e-06	1.08e-04		-6.60e-05	-3.05e-05
x (child's age at purchase)^2		(2.64e-05)	(1.09e-04)		(8.65e-05)	(3.67e-04)
(Parents' brand = child's brand)			-1.20e-06			-4.21e-07
x (child's age at purchase)^3			(1.34e-06)			(4.62e-06)
(Parents' brand = child's brand)	-2.94e-06	-3.03e-06	-3.39e-06	-2.80e-06	-1.79e-06	-1.90e-06
x months since purchase	(2.98e-06)	(3.04e-06)	(3.31e-06)	(5.29e-06)	(5.06e-06)	(4.80e-06)
Month of purchase fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Child's demographics	Yes	Yes	Yes	Yes	Yes	Yes
Parents' demographics	Yes	Yes	Yes	Yes	Yes	Yes
Child's state fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Parents' state fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Child's county fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Child's census tract fixed effects	No	No	No	No	No	No
Number of choice pairs	58,420	58,420	58,420	31,078	31,078	31,078
R^2	0.216	0.216	0.216	0.265	0.265	0.265

Standard errors clustered by 1968 PSID family are in parentheses. Each column is a linear probability model where each individual-year-vehicle choice enters once for each brand. Child's and parents' demographics include age, education, income, gender, number of children in household, and family size. All control variables are interacted with dummies for each brand.

have had more exposure to a car that their parents owned if they lived with their parents during the ownership period. This interacted variable may also be capturing age effects, however, because children are on average younger during the time in which they live with their parents. If children are more impressionable while they are young, then our interaction term may be positive because of a correlation with age. We address this possibility by allowing the relationship between the parents' brand choice and the child's brand choice to vary flexibly with the child's age at the time of the parents' purchase. Age effects may be nonlinear, so we include interactions between whether the parents' brand choice matches the child's brand choice and polynomials up to a cubic in the child's age. Finally, our regression specification also accounts for decay in brand preference transmission over time by controlling for the length of time between the child's and parents' purchases (interacted with the child's brand = parents' brand indicator).

Table 7 reports linear probability model regressions that add these interacted regressors in the expanded set of child-parent matched vehicles. We use our preferred specification with demographic controls and county fixed effects in all regressions. Columns 1 through 3 use all seven brands and include progressively richer polynomials in the child's age at the time of the parents' purchase; columns 4 through 6 repeat these specifications for the Ford and GM subsample. Our focus is on the interaction between the brand choices of parents and whether or not the child lived at home when the vehicle was purchased (the second variable in table 7). The estimated coefficient on this interaction term is positive in all specifications and is statistically significant in all three columns of the Ford and GM subsample, providing evidence that at least some of the intergenerational brand choice correlation is being driven by state dependence. Overall, this result is robust to increasing the richness of the age-at-purchase polynomial, and the polynomial coefficients themselves are consistent with intergenerational brand choice correlation being stronger for older children.[28] This age effect reflects the possibility that once they have moved out of the house, older children are more attuned to vehicle brands than younger children (perhaps due to budget constraints early in adult life). Finally, the interaction between the parents' brand choice dummy and the time elapsed between the parents' and child's purchases is negative (though insignificant) in all specifications. This last result suggests that the influence of the parents' choice on the child's choice decays over time, consistent with the findings from the lag models studied in section 4.3.[29]

5 Implications of brand preference transmission for the vehicle market

What might intergenerational brand preference transmission imply for firms' strategies and market outcomes in the automobile industry? Some implications are closely related to estab-

[28] Omitting age effects entirely reduces the estimate of the coefficient of interest by roughly one-third. This result occurs because the "at home" effect is conflated with the effect of purchases occurring at a young age.

[29] In investigating state dependence, we have also explored whether the intergenerational correlation in brand choice is weaker if the parents had a poor experience with a vehicle. To identify vehicles that were likely of poor quality, we used wholesale auction price data to estimate a depreciation rate for each vintage of each model and matched these depreciation rates to the parents' car choices. While our point estimates on the interaction of depreciation rates with parents' brand choice indicate that the intergenerational correlation is weaker for vehicles that proved to be of low quality, these regressions lack statistical power and we thus omit them from the paper. As an alternative approach, we simply interacted parents' brand choice with the length of time that the parents owned their vehicles; we imagine that parents would get rid of disappointing vehicles more quickly. While the intergenerational correlation is weaker for cars that parents held for shorter periods, this result does not necessarily indicate state dependence, for parents would also hold vehicles for shorter periods if the vehicles were a poor fit for the parents' underlying preferences.

lished findings in the literature regarding within-household brand attachment. For example, Bronnenberg et al. (2012) shows that state dependence in brand choice will strengthen incumbent firms by increasing barriers to entry and limiting the speed with which market shares may change over time. It is intuitive to expect that intergenerational preference transmission will exacerbate these effects by tying young consumers—who would otherwise be unattached to a brand and therefore open toward new entrants or smaller firms—to their parents' preferred brand.

Our findings also relate to the literature on switching costs, which focuses on how state dependence in consumers' brand choices affects firms' profits in equilibrium (Klemperer 1987; Dubé et al. 2009; Somaini and Einav Forthcoming). Incorporating intergenerational state dependence into traditional switching cost models yields unique implications in settings like the automobile industry in which firms offer multiple products targeted at consumers of different ages. To see the intuition, consider a simple overlapping generations model in which every consumer lives for two periods, buys two cars during his or her lifetime—one entry-level model while young and one upscale model while old—and has a child while old that becomes a young consumer in the next period. If consumers have state dependent brand preferences and do not pass these preferences on to their children, then this model resembles the one analyzed in Klemperer (1987), which establishes that firms will lower prices for young consumers to "invest" in brand loyalty and then "harvest" that loyalty among older consumers by charging higher prices.

Intergenerational state dependence disrupts this logic. In the extreme, if the dependence of children's preferences on their parents' choices is as strong as within-generation state dependence, then the overlapping generations become equivalent to an infinitely lived consumer with constant loyalty. In this case, the problem resembles the switching cost model of Dubé et al. (2009), in which firms balance harvesting and investing incentives in every period, and there is no economic distinction between young and old consumers. Thus, relative to the model with no intergenerational state dependence, equilibrium prices for upscale cars should fall, since firms will recognize that purchases by parents today improve their market share among young consumers tomorrow. Conversely, prices for entry-level cars should rise because young consumers enter the market with brand loyalty that can be harvested.

In appendix A, we set up and solve a simple version of this model with two symmetric Bertrand-competing firms that sell cars to consumers who live two periods each, buying a different type of car each period. Consumers have logit demand with a brand switching cost that we calibrate using our estimates from section 4.[30] For simplicity, we assume that the young and old car markets have the same cost and demand parameters. When we allow

[30]We also use within-household brand choice correlations that we discuss in appendix B.

for within-household but not intergenerational state dependence, the model reproduces the Klemperer (1987) style result that equilibrium prices are lower for cars targeted at young consumers than for cars targeted at older consumers. However, when intergenerational state dependence is as strong as within-household state dependence, the model is equivalent to that of Dubé et al. (2009), and equilibrium prices are equal for both types of cars. We also present an intermediate case in between these two extremes.

In our calibrated examples, the net effect of state dependence on average prices and firm profits is negative. This result echoes a similar finding by Dubé et al. (2009) in consumer packaged goods markets: when switching costs are modest, their presence may cause prices and profits to decline because firms are never able to fully capitalize on their customers' brand loyalty as they constantly compete for the next generation of consumers.[31] This finding may also speak to the auto industry's apparent focus on sales volumes to the potential neglect of current profits. The industry media is filled with stories about market share, sales volumes, and conquest rates. Anecdotally, automakers are said to focus on hitting quarterly sales targets, deeply discounting or even selling some vehicles at a loss to meet these targets. It is natural for an economist to view such prioritization of sales volumes over profitability as a mistake. In the presence of strong brand preferences, however, firms face a tradeoff between current and future profits, potentially justifying this focus on volume. Moreover, transmission of brand preferences across generations limits the incentive to harvest brand preferences among older consumers, as doing so might jeopardize the loyalty of future generations.

Intergenerational brand preference transmission may influence not only how firms set prices for the goods that they sell, but also the set of products they choose to develop. Most automakers offer a wide range of vertically differentiated products. State dependence provides one rationale for such a strategy. The stronger are brand preferences, the more valuable it is to keep consumers within the brand as they move through their life cycle and demand different types of cars.[32] When households have state-dependent brand preferences, producers have an incentive to develop entry-level offerings that will "lead" consumers to their profitable upscale goods as they age. If brand preferences are transmitted between generations, however, then producers of entry-level models will also have an incentive to develop upscale product lines to "lead" future generations to their entry-level products. The ability of upscale products to boost future sales of downmarket products may, for instance,

[31] Cabral (2009) further discusses why the investment incentive to lower prices is first-order while the harvesting incentive to raise prices is second-order.

[32] There are, of course, many other reasons that firms might offer a broad set of vertically differentiated goods apart from consumer brand preferences, including production economies of scope and the value of covering a wide range of consumers' attribute preferences.

help explain the relatively slow growth of Toyota and Honda in the 1980s, and Hyundai and Kia today, which offer mainly entry-level vehicles but do so at a low price given their high quality.

These considerations may also help explain the competition between Ford and GM early in the twentieth century. Ford initially sold a single, affordable vehicle and focused on driving down costs through economies of scale. Henry Ford had no interest in product differentiation and famously quipped that "people can have a Model T in any color—so long as it is black." Meanwhile, GM's strategy was to build a variety of cars to fit a range of lifestyles and income levels, embodied by the famous quote from Alfred Sloan that GM would sell "a car for every purse and purpose." As cars first became affordable to the masses, brand loyalty would have been minimal because most consumers had not previously owned a car, nor had their parents owned a car. This limited the initial benefits to GM's approach.

Over time, however, strong within-household brand preference would have allowed GM to charge higher prices to consumers who "graduated" to their upscale models. This would have given GM an extra incentive to cut prices on their entry-market cars, which were in competition with Ford's Model T, in order to gain future loyal upscale customers. Furthermore, having no upscale model would have also put Ford at a competitive disadvantage among the subsequent generation's entry-level consumers if intergenerational brand preference transmission was strong. That is, the children of consumers who had progressed to an upscale GM model would have inherited a preference for GM before they went to buy their first (entry-level) car. Thus, both within-household and intergenerational brand preference transmission may have been important in determining the ultimate success of GM's differentiated approach, which Ford itself later adopted.

6 Conclusion

Our analysis of PSID data suggests that automobile brand preferences may be passed through generations in ways that are important to the strategies of automobile producers. We document a strong correlation across generations in brand choice that remains strong even when limiting the analysis to similar brands and controlling for a rich set of demographic factors and fine geographic fixed effects, leading us to conclude that intrafamily correlations are likely not driven entirely by correlated demographic characteristics but rather reflect an important role for intergenerational brand preference transmission. Further, our finding that intergenerational choice correlation is stronger when children are directly exposed to their parents' vehicles suggests a role for intergenerational state dependence, in which actual experience with a vehicle is important for influencing brand choice across generations.

These results inform our understanding of endogenous preference formation, complementing recent work that has focused on the role of local tastes and geography in shaping consumers' preferences (Logan and Rhode 2010; Atkin Forthcoming; Bronnenberg et al. 2012). They may also inform automakers' pricing and product-line incentives. Intuition suggests, and our numerical simulations in the appendix confirm, that intergenerational state dependence curtails firms' ability to price discriminate across young and old consumers, since charging a high price to old consumers today reduces sales to young consumers tomorrow. That is, the "invest in young consumers and harvest old consumers" strategy (Klemperer 1987) is no longer optimal when parents' choices affect the preferences of their children. More broadly, intergenerational state dependence may also enhance firms' incentive to offer a broad range of products that appeal to consumers over their entire lifetime, thereby allowing multiple generations within a family to preserve their brand loyalty.

References

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, February 2005, *113* (1), 151–184.

Atkin, David, "Trade, Tastes and Nutrition in India," *American Economic Review*, Forthcoming.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, July 1995, *64* (4), 841–890.

Birch, Leann L., "Development of Food Preferences," *Annual Review of Nutrition*, 1999, *19*, 41–62.

Black, Sandra E. and Paul J. Devereux, "Recent Developments in Intergenerational Mobility," in "Handbook of Labor Economics," Vol. 4b, Elsevier, 2011, pp. 1487–1541.

Bronnenberg, Bart J., Jean-Pierre Dubé, and Matthew Gentzkow, "The Evolution of Brand Preferences: Evidence from Consumer Migration," *American Economic Review*, 2012, *102* (6), 2472–2508.

_ , Sanjay K. Dhar, and Jean-Pierre Dubé, "Brand History, Geography, and the Persistence of CPG Brand Shares," *Journal of Political Economy*, 2009, *117* (1), 87–115.

Cabral, Luís, "Small Switching Costs Lead to Lower Prices," *Journal of Marketing Research*, 2009, *46*, 449–51.

Doganoglu, Toker, "Switching Costs, Experience Goods, and Dynamic Price Competition," *Quantitative Marketing and Economics*, 2010, *8*, 167–205.

Dubé, Jean-Pierre, Gunter J. Hitsch, and Peter E. Rossi, "Do Switching Costs Make Markets Less Competitive?," *Journal of Marketing Research*, 2009, *46*, 435–45.

_ , _ , and _ , "State Dependence and Alternative Explanations for Consumer Inertia," *Rand Journal of Economics*, 2010, *41* (3), 417–45.

Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti, "Mothers and Sons: Preference Formation and Female Labor Force Dynamics," *Quarterly Journal of Economics*, 2004, *119* (4), 1249–1299.

Klemperer, Paul, "Markets with Consumer Switching Costs," *Quarterly Journal of Economics*, May 1987, pp. 375–394.

Langer, Ashley, "Demographic Preferences and Price Discrimination in New Vehicle Sales," Fall 2012. Manuscript: University of Arizona.

_ and Nathan Miller, "Automakers' Short-Run Responses to Changing Gasoline Prices," *Review of Economics and Statistics*, 2013, *95* (4), 1198–1211.

Logan, Trevon D. and Paul W. Rhode, "Moveable Feasts: A New Approach to Endogenizing Tastes," August 2010. Manuscript: The Ohio State University.

Luttmer, Erzo F.P. and Monica Singhal, "Culture, Context, and the Taste for Redistribution," *American Economic Journal: Economic Policy*, February 2011, *3* (1), 157–79.

Mannering, Fred and Clifford Winston, "A Dynamic Empirical Analysis of Household Vehicle Ownership and Utilization," *Rand Journal of Economics*, Summer 1985, *16* (2), 215–236.

_ and _ , "Brand Loyalty and the Decline of American Automobile Firms," *Brookings Papers on Economic Activity. Microeconomics*, 1991, pp. 67–114.

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

_ , "Economic Analysis of Social Interactions," *Journal of Economic Perspectives*, Summer 2000, *14* (3), 115–136.

Solon, Gary, "Intergenerational Income Mobility in the United States," *American Economic Review*, June 1992, *82* (3), 393–408.

—, "Intergenerational mobility in the labor market," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3 of *Handbook of Labor Economics*, Elsevier, October 1999, chapter 29, pp. 1761–1800.

Somaini, Paulo and Liran Einav, "A Model of Market Power in Customer Markets," *Journal of Industrial Economics*, Forthcoming.

Train, Kenneth E. and Clifford Winston, "Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers," *International Economic Review*, 2007, *48* (4), 1469–1496.

Online Appendix: Not for Publication

A Theory model details

This appendix provides additional information and results for the theory model that is described briefly in section 5. We focus on implications unique to intergenerational brand preference transmission by considering markets, such as those for automobiles, in which firms can discriminate between younger and older consumers. In particular, we explore how firms offering multiple products—each catering to a different age class—will want to price and market these products differently in the presence of intergenerational state dependence.[33]

The related literature (see especially Klemperer (1987), Dubé et al. (2009), and Somaini and Einav (Forthcoming)) has typically used the term "switching costs" to refer to a reduction in utility experienced by a consumer who switches from one brand to another in different periods. This is identical to how we implement what we here call state dependence.[34] Our model differs from the existing literature in that we allow the choices of older consumers to affect the preferences of their children (next period's younger consumers) and that we model firms as offering different products to consumers of different ages.[35] We model multi-product firms to relate the model more closely to the automobile market, in which nearly all manufacturers produce a range of models tailored to consumers in different stages of their lifecycle, and to highlight the role that intergenerational state dependence can play in determining automobile prices in equilibrium.

We study a simple, symmetric model in which two firms compete in two different product markets and consumers live two periods, purchasing once in each market. We forego

[33]As discussed in section 2 above, intergenerational brand preference transmission will only have implications for pricing strategies if it operates via the state dependence mechanism. In contrast, the direct preference inheritance mechanism will affect automakers' advertising and product line strategies but will not affect their pricing strategies.

[34]Some papers model switching costs as an increase in utility from purchasing the same brand that was purchased in the previous period (our approach), while others model switching costs as a decrease in utility from purchasing a different brand. Dubé et al. (2009) examines both models and finds that they produce identical predictions in the absence of an outside good. In the presence of an outside good, the second formulation yields lower prices in equilibrium, as switching costs make the outside good relatively more appealing.

[35]A further generalization of our model would be to allow for a broader set of peer effects so that, for example, older consumers' choices could influence other older consumers. While this generalization is beyond the scope of this paper, we believe that such peer effects would have similar effects to what we find here for intergenerational state dependence: a relative reduction in the markup for upscale vehicles (since within-household brand loyalty becomes less important in the presence of cross-household state dependence). Moreover, our focus on parent-to-child transmission is motivated by our empirical finding that the parent-to-child channel is stronger than that for other family links and by the extensive literature showing intergenerational correlations in many economic measures (Black and Devereux 2011).

a richer model that would more closely match the current automobile industry—a model with more than two major firms, many products per firm, and richly differentiated consumer preferences—for several reasons.[36] First, the simultaneous estimation of the parameters needed to simulate the model (those governing households' preference heterogeneity, households' brand preference transmission, and firms' marginal costs) would be a substantial undertaking that is beyond the scope of this paper and likely beyond the power of our data.[37] Second, the computational challenges of simulating such a model would be immense. Finally, the simple model we present is close in spirit to most of the brand loyalty literature and provides clear, intuitive results that we believe would generalize qualitatively to a more complex model.[38]

A.1 A simple model of automobile pricing under brand loyalty

In our model, there are two symmetric firms, denoted j and k, that compete in a differentiated Bertrand pricing game with an infinite horizon and overlapping generations of consumers. In each period, there are unit masses of two types of households: young (type A) and old (type B). All consumers are born as type A, become type B in the second period of their lives, and then die, creating a new type A consumer (offspring) upon death. All consumers purchase exactly one vehicle in each period of their lives, and there is no outside good. We have in mind that children are present in the type B households, are exposed to their parents' brand choice, and then become type A consumers upon leaving their parents' home. A key feature of the model is that the type A and type B consumers purchase different kinds of cars. Both firms are aware of this fact, and both sell two vehicle models catering to the two types. Thus, there are four vehicles in the market: jA, jB, kA, and kB. Car types A and B can be thought of as cars preferred by younger versus older consumers, or entry level versus upscale, or single-person versus family vehicles.

For both brevity and clarity, we will focus on the case in which type A households consider only vehicles jA and kA and type B households consider only vehicles jB and kB. Clearly

[36]The treatment of consumers as living for two periods is also an abstraction, though one that is in line with much of the switching cost literature (Klemperer (1987) and Somaini and Einav (Forthcoming), for example). We revisit the question of the duration of consumers' lifetimes and the time gap between periods when we discuss the model's discount factor further below.

[37]Dubé et al. (2009) and Dubé et al. (2010) are able to simultaneously estimate preference heterogeneity and within-household brand loyalty because they observe both a large number of repeat purchases per customer and rich price variation in their dataset on orange juice and margarine purchases. While our PSID dataset is well-suited for estimating intergenerational brand preference transmission, the limited number of purchases observed for each household and weak price data make it poorly suited for characterizing heterogeneous preferences for price and other attributes.

[38]We are encouraged here by the fact that Dubé et al. (2009) find qualitatively similar predictions from the simple and complex versions of their model.

this is an abstraction, as there will be some substitution by households across vehicle types.[39] Still, in a survey of over 22,000 consumers by a market research firm described in Langer (2012), the Cadillac Deville and Lincoln Town Car had more than 100 purchasers over the age of 60 and none under the age of 40, while the Scion tC had more than 100 purchasers under 40 and only 6 over 60. Similarly, only 5% of consumers who say they purchased a Buick are under the age of 40. Clearly, there are vehicles that appeal strongly to specific age groups.

Let the utility of a particular consumer i of type B that purchases vehicle jB be given by:

$$U_{ijB} = V - \alpha P_{jB} + \mu_B 1\{b_{iA} = j\} + \varepsilon_{ijB},$$

where V is a baseline utility that is common across the two brands, P_{jB} is the price of vehicle jB, and $1\{b_{iA} = j\}$ is an indicator for whether consumer i purchased brand j when he or she was a type A last period. The parameter μ_B denotes the strength of within-consumer persistence of brand preferences. The utility from purchasing the other brand's vehicle kB is given similarly.

The utility of a consumer i of type A that purchases vehicle jA is similarly given by:

$$U_{ijA} = V - \alpha P_{jA} + \mu_A 1\{b_{iB} = j\} + \varepsilon_{ijA}.$$

Here, $1\{b_{iB} = j\}$ is an indicator for whether the parents of consumer i purchased brand j when the parents were type B last period. The parameter μ_A denotes the strength of intergenerational brand preferences. This formulation assumes that the parents' type A car—which we imagine to be owned by parents before the next generation is born—does not influence the child's utility function. (Thus, our two-period formulation does not distinguish between short- and long-run state dependence, but it does exclude direct brand preference inheritance that could cause parental choices that occurred before a child was born to influence child choice.) The random utility components ε_{ijB} and ε_{ijA} are assumed to be i.i.d. type I extreme value over individuals i, brands j and k, and types A and B.

[39] Allowing for some cross-age substitution has essentially no impact on models in which intergenerational brand preference transmission is as strong as within-household transmission. In models in which intergenerational transmission is relatively weak, cross-age substitution reduces the gap between the type A and type B vehicle prices (and does so in a qualitatively symmetric way). The result that brand preferences (of a magnitude corresponding to our estimates above) reduce equilibrium prices continues to hold. This is true even in the extreme case in which there is no intergenerational brand preference and consumers have no systematic preference for their own type of vehicle. This last model is similar to that of Doganoglu (2010), in which consumers live for two periods and the (single product) firms cannot distinguish between young and old.

We assume that type A consumers are not forward-looking when deciding whether to purchase vehicle jA or kA.[40] We also assume that type B consumers are not forward looking in the sense that they do not consider the implications of the brand preferences they transmit to their children.

We next discuss aggregate demand and the firms' profit and value functions. Let ϕ_A and ϕ_B denote the fraction of consumers loyal to brand j in the A and B markets, respectively. Given the price of each vehicle and ϕ_A and ϕ_B, the demand for each vehicle will be given by a weighted sum of standard logit choice probabilities. For example, the demand for vehicle jA is given by:

$$\begin{aligned} D_{jA} = \phi_A & \frac{\exp(V - \alpha P_{jA} + \mu_A)}{\exp(V - \alpha P_{jA} + \mu_A) + \exp(V - \alpha P_{kA})} \\ & + (1 - \phi_A) \frac{\exp(V - \alpha P_{jA})}{\exp(V - \alpha P_{jA}) + \exp(V - \alpha P_{kA} + \mu_A)}. \end{aligned}$$

We model the marginal cost of all four vehicles in the market as a constant, denoted by c. Firm j's per-period profits are then given by:

$$\pi_j(P_{jA}, P_{kA}, P_{jB}, P_{kB}, \phi_A, \phi_B) = (P_{jA} - c) \cdot D_{jA}(P_{jA}, P_{kA}, \phi_A) + (P_{jB} - c) \cdot D_{jB}(P_{jB}, P_{kB}, \phi_B).$$

In the infinitely repeated game, the firms' state variables are the brand loyalty shares ϕ_A and ϕ_B of the consumers of each type. The states evolve so that next period's loyalty of the type A consumers is given by the current period's demand of the type B consumers for vehicle jB: $\phi_A^{'} = D_{jB}(P_{jB}, P_{kB}, \phi_B)$. Similarly, $\phi_B^{'} = D_{jA}(P_{jA}, P_{kA}, \phi_A)$. We restrict the firms to Markov strategies so that, with a discount factor δ that is shared by the two firms, firm j's Bellman equation is given by:

$$V_j(\phi_A, \phi_B) = \max_{P_{jA}, P_{jB}} \{\pi_j(P_{jA}, P_{kA}, P_{jB}, P_{kB}, \phi_A, \phi_B) + \delta V_j(\phi_A^{'}, \phi_B^{'})\}$$

Firm k's Bellman equation is defined similarly. These equations capture the tradeoff the firms face as the parameters μ_A and μ_B, which govern the strength of brand loyalty, vary. The incentive to increase current-period profits by increasing prices is weighed against the incentive to increase future profits by lowering prices to boost the share of future loyal consumers.

For a given set of model parameters, the Markov Perfect Equilibrium (MPE) of the firms'

[40] Per the intuition of Somaini and Einav (Forthcoming), we expect that allowing for forward-looking behavior by type A consumers would result in higher prices for type A vehicles because type A consumers will become less sensitive to current price changes. As a second-order effect, prices for type B vehicles should then fall in equilibrium because the continuation value of future type A consumers will have increased.

dynamic Bertrand pricing game can be solved computationally using value function iteration techniques.[41] In the simulations presented below, we fix $V = 1$, $\alpha = 8$, and $c = 1$. The choice of V is immaterial in the absence of an outside good. The price preference α and marginal cost c parameters together yield, in the absence of any brand preferences, an equilibrium price for all vehicles of 1.25 and equilibrium own-price elasticities of -5. This markup and elasticity roughly correspond to typical markups and elasticities found by Berry, Levinsohn and Pakes (1995).

The choice of discount factor merits discussion. We treat automakers as having an annual real discount factor of 0.9 and treat the time between periods in the model as five years, which roughly corresponds to the average vehicle holding time in our data. Thus, the discount factor δ used in our model is $0.9^5 \approx 0.59$. An obvious tension here is that consumers live longer than ten years. One alternative approach would be to use a discount factor that reflects the time gap between generations (with a value of $0.9^{25} \approx 0.07$, for example). In this case, automakers would care little about future generations when setting prices. However, this alternative approach neglects the fact that consumers purchase vehicles more frequently than every 25 years and that children begin purchasing vehicles of their own soon after they are exposed to purchases their parents made while they were teenagers still living at home.

Ideally, we would resolve this issue by studying a model in which consumers live for many periods, each five years apart, and transfer brand preferences to new consumers (their children) via their brand choices over one or several purchases relatively late in life. In such a model, even though generations would be far apart in time, the chain of purchases occurring every five years—and in particular across the short, potentially overlapping transition from one generation to the next—would give automakers a sufficient incentive to consider the next generation when setting prices for upscale vehicles. While the high dimensionality of the state and decision spaces of such a model precludes its implementation, a previous version of this paper did present a computationally feasible version of such a model that forces automakers to group consumers into two broad classes (young and old), selling one vehicle type to each class and forbidding within-class price discrimination. Despite this alternative model's large time gap between generations, we nonetheless found qualitatively similar results to those discussed below. In particular, intergenerational state dependence eliminates differences in markups between vehicles targeted at young versus old consumers.

Finally, the range of brand loyalty parameters μ_A and μ_B that we consider spans zero to one. Values of zero collapse the model to a standard static Bertrand problem, for which the

[41] Without intergenerational brand loyalty ($\mu_A = 0$), the model reverts to a standard two-period game (akin to that of Klemperer (1987)) that can be characterized analytically, though the results presented below for this case were nonetheless generated numerically.

equilibrium price is 1.25. Our estimates from section 4 correspond to values of μ_A within the zero to one range. For our preferred "all brands" specification (column 5 of table 2), the corresponding value of μ_A is about 0.35.[42] For the Ford/GM regression, μ_A is about 0.17, while μ_A is about 0.65 for the Toyota/Honda regressions. To calibrate values of μ_B, we have estimated within-household brand choice correlations that parallel the intergenerational correlations presented in tables 2 and 3. These regressions, which are discussed in more detail in appendix B, suggest that μ_B is about 0.75 in the regression with all brands and 0.50 in the Ford/GM regressions—so roughly 2-3 times as large as μ_A.

A.2 Optimal prices in a model with symmetric firms

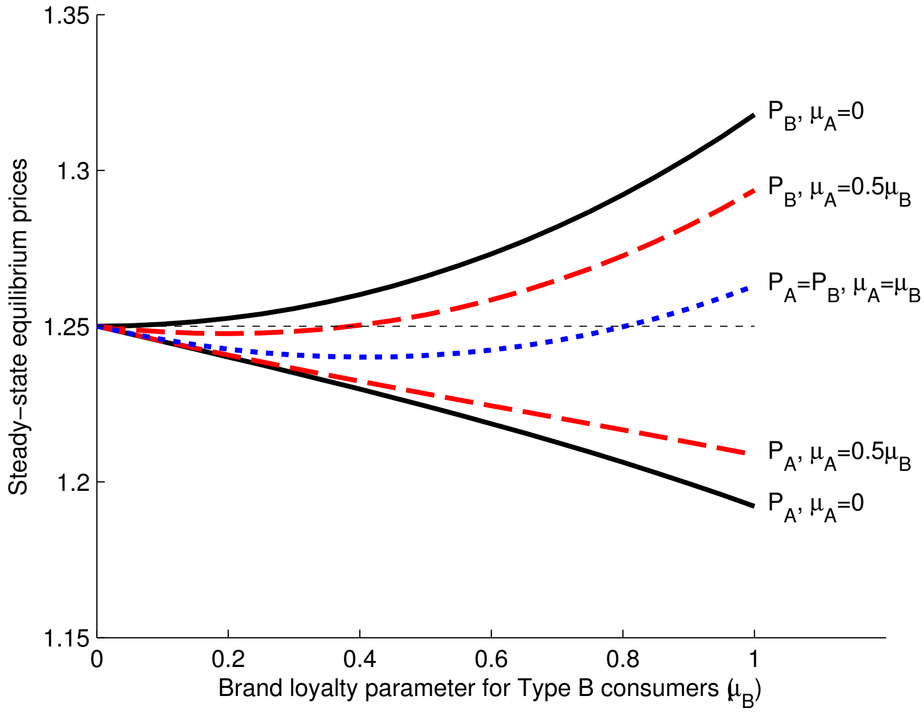
We explore the impact of brand preferences on firms' equilibrium pricing strategies by increasing the brand preference parameters μ_A and μ_B from zero and examining the change in firms' equilibrium steady state prices. These prices are sufficient statistics for steady state profits because in steady state the two firms split the A and B markets equally (due to the symmetry of the firms' demand and cost parameters).

Figure 1 presents steady state equilibrium prices, over a range of brand loyalty strengths, for three cases. For all cases, the prices of firms j and k are equal within each of the markets A and B due to symmetry. In the first case, given by the solid lines, intergenerational brand transmission is turned off by holding $\mu_A = 0$ while the strength of within-household brand preference is varied by letting μ_B range from 0 to 1. In this case, we find that increasing μ_B raises the prices of the type B cars while lowering the prices of the type A cars. That is, when households develop brand loyalty but do not pass this loyalty to their children, the equilibrium prices for vehicles intended for older consumers will be high relative to prices for vehicles intended for younger consumers. The intuition for this result follows directly from Klemperer (1987): if first period choices determine brand loyalty in the second period, then firms will "invest" in customers in the first period by charging lower prices and "harvest" the consumer loyalty in the second period. The "investment" effect in the A market outweighs the "harvesting" effect in the B market (that is, average vehicle price is less than the no-loyalty baseline price of 1.25) for values of μ_B up to about 0.83. If brand loyalty is stronger than that, however, then the "harvesting" effect dominates in our model.

When intergenerational brand loyalty is equal to within-household brand loyalty—the case denoted by the dotted line in figure 1—the A and B markets behave identically to one another so that the prices for all four vehicles are equal in steady state, and the model

[42] Given a value for μ_A and assuming the brands in the choice set yield equivalent utility in the absence of a brand preference, the effect of parents' ownership on the probability of brand choice in our model is given by $(e^{\mu_A} - 1)/(e^{\mu_A} + (N-1))$, where N is the number of brands in the choice set.

Figure 1: Steady state prices with two symmetric firms



Note: Steady state equilibrium prices shown are from the model described in section A.1 in which $\delta = 0.9^5$, $V = 1$, $\alpha = 8$, and $c = 1$. At steady state, the demand for each of the four cars jA, jB, kA, and kB is equal to 0.5. The solid line denotes the case in which there is no intergenerational brand loyalty, the dashed line denotes the case in which intergenerational brand loyalty is half the strength of within-household brand loyalty, and the dotted line denotes the case in which intergenerational and within-household brand loyalty are equal.

collapses to that of Dubé et al. (2009). Relative to the case with no intergenerational state dependence, type B prices fall because high type B prices now reduce future demand and profits, and type A prices rise because investing in future type B consumers is no longer as profitable. The prices of the type A and B vehicles—now equal in steady state—are roughly equal to the average of the type A and B prices from the no intergenerational state dependence case. That is, intergenerational state dependence appears to primarily affect the distribution of prices across types rather than the average price in the market. Thus, similar to the no intergenerational state dependence case, steady state equilibrium prices are lower than in the case of no brand loyalty for values of μ_B up to about 0.80.

Finally, the dashed line plots an intermediate case in which intergenerational brand preference parameter, μ_A, is half as large as the within-household parameter, μ_B. This case is consistent with our empirical estimates of the relative strength of intergenerational state dependence and within-household state dependence. Not surprisingly, this case lies between the two other cases. Here, average vehicle price is less than the no-loyalty baseline price of 1.25 for values of μ_B up to about 0.97, which exceeds our preferred estimates for μ_B of about 0.5 to 0.75. This implies that the existence of brand loyalty causes a net reduction in firm profits, which accords with the theoretical intuition of Cabral (2009).[43]

B Within-household brand loyalty

In order to understand the size of intragenerational brand loyalty relative to within-household brand loyalty, we compute the size of within-household brand loyalty in our data. To that end, we use the sample of households whose purchases can be matched to their parents' prior vehicle purchase and estimate regressions analogous to those in table 2 of the text, with the brand of the household's most recent purchase in place of the brand of the parents' most recent purchase. Columns 1 through 4 of Table 8 present these results. We do not include specifications that include census tract fixed effects because these present an incidental parameters problem in what is essentially a lagged dependent variable regression.

Table 8: Correlations between household brand choice and previous household brand choice

	Dependent Variable: Household's Brand			
VARIABLES	(1)	(2)	(3)	(4)
Household's brand = Lagged Household brand	0.225	0.218	0.203	0.137
	(0.007)	(0.007)	(0.007)	(0.007)
Month of purchase fixed effects	Yes	Yes	Yes	Yes
Household's demographics	No	Yes	Yes	Yes
Child's state fixed effects	No	No	Yes	Yes
Household's county fixed effects	No	No	No	Yes
Household's census tract fixed effects	No	No	No	No
Number of choices	17,268	17,268	17,268	17,268
R^2	0.117	0.124	0.138	0.220

Standard errors clustered by 1968 PSID family are in parentheses. Each column is a linear probability model where each individual-year-vehicle choice enters the data 7 times, once for each brand (GM, Ford, Chrysler, Toyota, Honda, Other Asian, and European). Household's demographics include age, education, income, gender, number of children in household, and family size. All control variables are interacted with 7 dummies, one for each brand. Columns 5 and 6 limit the sample to households living in census tracts that contain more than one PSID family.

We then run the same specifications using the subsample of households that purchased a Ford or GM vehicle. As in the intergenerational brand preference results, we do not require

[43]As noted above, a previous version of this paper contained a more detailed model in which consumers purchased cars multiple times while young and old. Results from this more detailed model also support this conclusion.

that the household's previous vehicle purchase was also a Ford or GM, but we do include a dummy variable that is equal to 1 if the previous vehicle was a Ford or GM. Table 9 shows within-household brand loyalty results that are analogous to the intergenerational brand loyalty results presented in table 3 of the text.

Table 9: Correlations between household vehicle brand and previous household brand choice among those owning a Ford or GM

	Dependent Variable: Household's Brand			
VARIABLES	(1)	(2)	(3)	(4)
Household's brand = Lagged Household brand	0.369	0.368	0.345	0.243
	(0.019)	(0.019)	(0.019)	(0.020)
Month of purchase fixed effects	Yes	Yes	Yes	Yes
Household's demographics	No	Yes	Yes	Yes
Household's state fixed effects	No	No	Yes	Yes
Household's county fixed effects	No	No	No	Yes
Household's census tract fixed effects	No	No	No	No
Number of choices	9,355	9,355	9,355	9,355
R^2	0.136	0.137	0.153	0.277

Standard errors clustered by 1968 PSID family are in parentheses. Sample is limited to the cases where the child chose Ford or GM. Each column is a linear probability model where each individual-year-vehicle choice enters the data 7 times, once for each brand (GM, Ford, Chrysler, Toyota, Honda, Other Asian, and European). Household's demographics include age, education, income, gender, number of children in household, and family size. All control variables are interacted with 7 dummies, one for each brand. Columns 5 and 6 limit the sample to households living in census tracts that contain more than one PSID family.