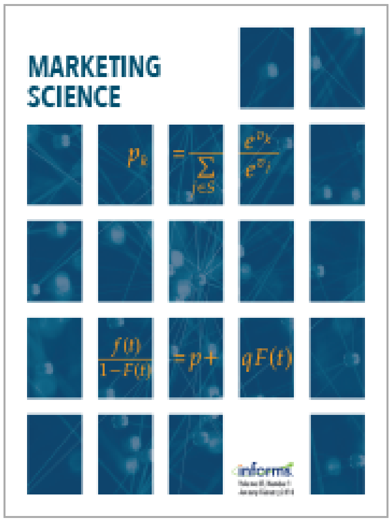


Network Effects in Alternative Fuel Adoption: Empirical Analysis of the Market for Ethanol

Scott K. Shriver

MARKETING SCIENCE
Vol. 34, No. 1, January–February 2015, pp. 78–97
ISSN 0732-2399 (print) | ISSN 1526-548X (online)
http://dx.doi.org/10.1287/mksc.2014.0881

Network Effects in Alternative Fuel Adoption: Empirical Analysis of the Market for Ethanol

Scott K. Shriver
Columbia Business School, New York, New York 10027, ss4127@columbia.edu

This paper investigates the importance of network effects in the demand for ethanol-compatible vehicles and the supply of ethanol fuel. An indirect network effect, or positive feedback loop, arises in this context due to spatially-dependent complementarities in the availability of ethanol fuel and the installed base of ethanol-compatible vehicles. Marketers and social planners are interested in whether these effects exist, and if so, how policy might accelerate adoption of the ethanol fuel standard within a targeted population. To measure these feedback effects, I develop an econometric framework that considers the simultaneous determination of ethanol-compatible vehicle demand and ethanol fuel supply in local markets. The demand-side model considers the automobile purchase decisions of consumers and fleet operators; the supply-side model considers the ethanol market entry decisions of competing fuel retailers. The framework extends extant market entry models by endogenizing the market size shifting fuel retailer profits. I estimate the model using zip code panel data from four states over a nine-year period. The model estimates provide evidence of a network effect. Under typical market conditions, entry of an additional ethanol fuel retailer leads to a 6% increase in the probability of ethanol-compatible vehicle purchase. The entry model estimates imply that the first entrant requires a local installed base of approximately 300 ethanol-compatible vehicles to be profitable. As an application, I demonstrate that subsidizing fuel retailers to offer ethanol in selective geographic markets can be an effective policy to indirectly increase ethanol-compatible vehicle sales.

Keywords: indirect network effects; market entry; alternative fuels; ethanol; flex-fuel vehicles
History: Received: September 22, 2012; accepted: August 4, 2014; Preyas Desai served as the editor-in-chief and Jean-Pierre Dubé served as associate editor for this article.

1. Introduction

In this paper, I examine the role of network effects in the market for ethanol as a transportation fuel. Although ethanol is a common additive to most forms of motor fuel, my inquiry relates specifically to the adoption of E85, a blend containing 85% ethanol and 15% gasoline. E85 is classified as an alternative fuel that may only be used in flex-fuel vehicles (FFVs) that are engineered to process fuel blends of up to 85% ethanol. As with classical studies of hardware/software systems (e.g., Chou and Shy 1990; Church and Gandal 1992, 1993; Katz and Shapiro 1994), the network effects of interest here are indirect because they operate through a system of complementary goods. That is, positive feedback arises in this context because as the installed base of hardware (i.e., FFVs) grows, more firms enter the market for compatible software (i.e., E85), which in turn leads to more hardware sales, and so on. The fact that FFV demand and E85 market entry are endogenously determined makes identification of network effects a challenging empirical exercise. My first objective for the study is therefore to consistently measure each side of the feedback loop, i.e., the effect of E85 availability on FFV demand, and the effect of the FFV installed base on the supply of E85. My second objective is to explore the marketing policy implications of these measurements for suppliers of FFVs and E85 fuel. A third and more general objective of the paper is to advance an empirical strategy permitting inference of simultaneous supply and demand in markets with endogenous market sizes, such as those governed by network effects.

To achieve these goals, I develop an equilibrium model of FFV demand and E85 market entry. The model may be interpreted as a two-sided technology adoption process, with observations of FFV acquisition (E85 entry) representing vehicle operator (fuel retailer) decisions to adopt the ethanol fuel standard.[1] I model FFV demand as a discrete choice of whether to purchase an E85-compatible vehicle, which is a function of the number of E85 retailers in the local market. The number of E85 retailers is in turn modeled as the outcome of a competitive market entry game of complete information, in which the installed

[1] In the terminology of the technology standards literature (e.g., Farrell and Saloner 1985), E85 defines a standard of interoperability for vehicles and fuels, in the same manner that the DVD format defines an interoperability standard for media players and titles.

base of FFVs determines the size of the market in which E85 retailers compete. To estimate the model, I assemble a rich panel data set that encompasses the entire population of FFV purchases and E85 market entry events in roughly 4,000 Midwestern zip codes over the nine-year period 2001–2009.

The panel data format provides several key advantages for econometric identification. First, zip code level spatial resolution facilitates direct observation of the localized markets in which the network effects of interest operate. Second, repeated observation of these markets admits extremely strong controls for unobservables, mitigating concerns over a leading source of endogeneity. Third, under the assumption that market outcomes are sequentially exogenous, lagged endogenous regressors become available as instruments to further address endogeneity concerns arising from the simultaneous determination of FFV demand and E85 supply. I exploit these features in a two-step estimation procedure. In the first step, I estimate the FFV demand parameters using the System GMM (SGMM) estimator of Blundell and Bond (1998), instrumenting for the number of E85 retailers in the market. In the second step, I condition on the FFV demand estimates and estimate the E85 market entry model by minimizing differences between observed and model-predicted entry. Entry predictions are generated via a novel simulation-based method that solves the reduced form of the FFV demand/E85 entry system for the equilibrium number of E85 retailers, thereby accounting for the endogeneity of FFV demand when estimating the E85 retailer profit function.

The model estimates provide evidence of a network effect. Consumer FFV demand estimates suggest that, under average market conditions, entry of an additional E85 retailer leads to a 6% increase in the probability of FFV purchase. I find a comparably-sized effect of E85 availability on fleet FFV demand.[2] Market entry model estimates imply that an E85 retailer can operate profitably with an installed base of approximately 300 consumer FFVs or an equivalent mixture of consumer and fleet FFVs, where fleet FFVs deliver 1.4 times the profit contribution of consumer FFVs. To demonstrate the application of the model estimates, I: (a) quantify the long-run impact of the network effect on observed FFV sales and E85 market entry, and (b) assess the benefits and costs for FFV manufacturers to subsidize the entry of E85 retailers. The first simulation attributes 5.6% of historical FFV purchases and 5.3% of E85 market entry events to the network effect. In the second application, I show that subsidizing E85 market entry can accelerate the adoption of FFVs, but that the cost effectiveness of this strategy depends critically on targeting promising local markets for intervention.

The paper contributes to several streams of research, the most central of which is the empirical literature on indirect network effects. Most studies of indirect network effects have considered high tech products such as consumer electronics and related media titles. Among the products considered by these studies are VCRs (Ohashi 2003, Park 2004), CD/DVD players (Basu et al. 2003, Dranove and Gandal 2003, Gandal et al. 2000, Karaca-Mandic 2011), personal digital assistants (PDAs) (Nair et al. 2004), and video game consoles (Clements and Ohashi 2005, Dubé et al. 2010, Lee 2013, Liu 2010, Shankar and Bayus 2003). As high tech products are distributed through well established brick-and-mortar and e-commerce channels, feedback in these markets operates by increasing the *variety* of compatible software rather than by expanding its spatial *availability*, as studied here. Whereas the previous literature has modeled the aggregate supply of software as a reduced form under free entry or monopolistic competition, accounting for the build-out of new distribution infrastructure in my setting calls for an oligopolostic model of fuel retailer entry into local markets. Empirical measurement of such spatial network effects has been the subject of limited research to date. One notable exception is Rysman (2004), who investigates indirect network effects in the market for Yellow Pages directories, which are spatially dependent by nature on their circulation areas. However, Rysman (2004) does not explicitly consider firm market entry decisions, focusing instead on the role of network effects for competition among incumbent firms. A separate literature on shopping malls (e.g., Eppli and Benjamin 1994, Vitorino 2012) examines spatial network effects arising through the agglomeration of demand rather than the diffusion of supply.

The paper also adds to the literature on empirical models of market entry. My model follows the approach of Bresnahan and Reiss (1990, 1991, 1994) to infer retailer profit functions from observations of the number of firms in the market. This class of models traditionally uses exogenous variation in market size to separately identify variable profits and fixed costs. I extend the Bresnahan and Reiss framework to allow for an endogenously determined market size, as is implied by a network effect or other situations where entry affects market potential.[3] Also related in

[2] Fleet vehicles are owned and operated by corporations or government agencies, as opposed to individual consumers.

[3] A related approach is outlined in Berry and Waldfogel (1999), but their model is more data-intensive, requiring observations of "software" sales and prices to estimate firms' variable profits. My model is more suitable to contexts where the entire profit function must be inferred from observations of firm market entry, in combination with rich data on consumer "hardware" adoption.

this sense are new models in the product-choice literature that combine demand data with discrete choice models of firms' location choices in product attribute space (e.g., Draganska et al. 2009). Relative to this literature, my contribution accommodates endogeneity concerns that arise when entry is potentially conditioned on demand-side unobservables.

From a policy perspective, the paper also contributes to recent research on the economics of ethanol as a transportation fuel in the United States. Two recent studies include Anderson (2011), who estimates demand for E85 and finds evidence of fuel-switching behavior among FFV owners, and Anderson and Sallee (2011) whose findings suggest that domestic automakers produce FFVs as a means to lower the cost of complying with federal fuel-economy Corporate Average Fuel Economy (CAFE) standards.[4] In a closely related paper, Corts (2010) investigates the mandated use of FFVs by government fleets on the build-out of E85 distribution infrastructure. Corts (2010) develops a reduced form regression model of the number of E85 retailers, but does not explicitly consider FFV demand, precluding counterfactual experiments of the sort performed here. As in my study, Corts (2010) obtains positive effects of the flex-fuel installed base on the number of E85 retailers and finds that fleet FFVs contribute more to E85 entry than do privately owned FFVs. I compare the findings in more detail when presenting my results in §6.2.

The paper proceeds as follows. In §2 I provide an overview of the ethanol fuel industry, followed by a description of my data in §3. I develop the econometric model in §4 and discuss estimation and identification issues in §5. Section 6 presents the main estimation results. In §7, I present applications of the results and discuss their strategic implications for firms. Section 8 concludes with a summary and a discussion of directions for future research.

2. Ethanol Fuel Industry

Ethanol's use as a transportation fuel has grown dramatically over the past decade, reaching 8% of U.S. fuel consumption by volume in 2009 (U.S. EIA 2011). The overwhelming majority of ethanol in the U.S. is consumed as an additive to "regular" gasoline, which commonly contains up to 10% ethanol (E10). The blending of ethanol into gasoline has been spurred by a series of public policy initiatives aimed at improving air quality and increasing the use of renewable fuels. Adding ethanol to gasoline increases the oxygen content of the blend, which in turn reduces air pollution from fuel combustion. Many states have mandated that ethanol replace methyl tertiary-butyl ether (MTBE) as an oxygenate, due to concerns over MTBE toxicity. The major federal incentive encouraging the production of ethanol is the renewable fuel standard (RFS), which requires that fuel suppliers increase the use of domestically produced biofuels to 36 billion gallons by 2022.

Blends of 85% or more of alcohol, including E85, are classified as alternative fuels that qualify for additional federal and state incentives. The use of E85 has grown steadily since the introduction of FFVs in 1998, and is particularly concentrated in the Midwest. For example, in Minnesota, the leading state in E85 adoption, E85 was available at 13% of the state's service stations in 2009 (compared to less than 1% nationally) and comprised 5% of statewide fuel sales. The concentration of E85 in the Midwest is linked to its abundance of corn, the primary feedstock for ethanol in the United States. Ethanol refineries are built close to corn growing regions to minimize feedstock transportation costs. Furthermore, finished ethanol is expensive to move long distances because it cannot be transported via existing oil pipelines (due to its corrosive and water-absorbing properties) and instead must be moved by truck, railcar, or barge.

Commercial production of FFVs that use E85 is dominated by domestic U.S. manufacturers, who tend to supply flex-fuel versions of their top selling trucks, SUVs, sedans, and minivans. As of 2009, 36 different FFV models were produced, constituting 14% of all vehicles supplied to the U.S. market (U.S. EIA 2012). Apart from differences in fuel economy, FFVs are identical to their gasoline-only counterparts with respect to operation, performance, and maintenance. According to EPA data (www.fueleconomy.gov) for model years 2001–2009, compared to gas-only variants, FFVs averaged 26% fewer MPG when using E85 (due to ethanol's lower energy density than gasoline) and 2% fewer MPG when using regular gasoline (due to suboptimal injection timing). Limited modifications are required to create flex-fuel variants of existing vehicles. For example, fuel storage tanks and transfer lines must be upgraded to prevent corrosion and fuel injection systems must be altered to detect and adjust to varying fuel composition. The cost for manufacturers to provide these upgrades is relatively modest, generally less than $200 per vehicle. Manufacturers typically do not pass through these additional costs in their suggested retail prices (MSRPs), a policy consistent with Anderson and Sallee's view

[4] CAFE standards require vehicle manufacturers to maintain an overall fuel efficiency rating for the fleet of vehicles they supply. Under the Alternative Motor Fuels Act of 1988, vehicle manufacturers are granted CAFE credits for producing alternative fuel vehicles (including FFVs), offsetting lower fuel economy in the remainder of their product line. However, limitations are placed on the extent to which this compliance loophole may be used, i.e., FFV production can reduce the fuel economy rating for a manufacturer's product line by a maximum of 1.2 miles per gallon (MPG).

that FFV production is motivated, at least in part, by reducing CAFE regulation compliance costs through accumulating alternative fuel vehicle (AFV) credits. I consider the implications of such manufacturer policies for the measurement of network effects in my discussion of model identification in §5.2.

2.1. Vehicle Operator Adoption of FFVs

Two distinct groups potentially operate FFVs and use E85: consumers and corporate/government fleets. Although my primary interest is measuring feedback in consumer FFV acquisition and E85 market entry, fleets play an important role in this market because of strong government incentives. Specifically, the Energy Policy Act of 1992 requires that certain fleets maintain a minimum proportion of AFVs.[5] Various state statutes also mandate fleet AFV acquisition and alternative fuel use whenever available and economically feasible. To account for such systematic differences in alternative fuel adoption, I model consumer and fleet populations separately in the analysis that follows.

Absent alternative fuel adoption mandates, consumers choose whether to acquire an FFV by weighing the perceived benefits against any perceived costs or risks. A prerequisite for an informed decision is awareness of E85 and the dual fueling capability of FFVs. Surveys indicate that such awareness is common in regions where E85 is generally available. For example, Phoenix Automotive (2006, 2007) found 67% awareness of E85 across the Midwest and 100% awareness of FFV E85 compatibility among FFV owners in Minnesota.[6] Because FFVs are backwardly compatible with gasoline, the primary benefit to consumers is the *option* to use E85. Surveys indicate that consumers value the E85 fueling option on the basis of: (a) potential fuel cost savings, (b) reduced environmental impact (less air pollution, use of renewable resources), and (c) energy security/support for domestic industries (Harris Interactive 2006; Maritz Automotive Research Group 2006; Phoenix Automotive 2006, 2007).[7] Inasmuch as realization of utility from the aforementioned benefits is linked to the use of E85, consumer valuation of the E85 FFV fueling option should be increasing in the availability of E85. Conversely, in markets where E85 is scarce or unavailable, adverse factors such as risk aversion to nontraditional engines, lower fuel economy when using gasoline or limited selection of FFV models may dominate the dual fuel option value.

From a modeling perspective, the fact that the E85 fueling option is provided at little or no cost is convenient, as it greatly diminishes the scope for dynamic considerations in the FFV acquisition decision. That is, FFV purchase does not entail any (incremental) sunk acquisition costs or an irreversible commitment to using E85 fuel. Drawing on the logic of Dixit and Pindyck (1994), there is no reason to strategically defer a decision that is costlessly reversible because there is no meaningful intertemporal trade-off between investment costs and the accrued stream of benefits. In such circumstances, the optimal (dynamic) decision that maximizes the expected present value of future payoffs collapses to the optimal (static) decision conditioned on current information.

2.2. Fuel Retailer Adoption of E85

E85 is typically distributed by traditional fuel retailers, i.e., gas stations and convenience stores. However, government agencies and companies occasionally operate private fueling facilities that exclusively service E85 compatible vehicles in their fleet. As these private-access fueling facilities do not contribute to the indirect network effects of interest here (the benefits of E85 availability are internalized by the fleet operator and unrealized by other agents) and because they represent a fraction of E85 outlets ($<4\%$ of all outlets in my data, as of 2009), they play a limited role in the analysis that follows. That is, I do not explicitly model the fleet operator's decision to build a private fueling facility. Rather, my empirical strategy will be to control for the impact of private E85 facilities on fleet FFV demand and traditional fuel retailer entry into the E85 market, and to use instrumental variables to address potential endogeneity concerns.

For traditional fuel retailers, the profit contribution from offering E85 is contingent on several factors, i.e., competition, market size, variable costs, fuel usage, and willingness to pay for E85. The market size is measured by the local installed base of FFVs, and thus it is through the market size dependence that the network effect operates on the E85 supply side of the market. Variable E85 costs will be determined

[5] The specific terms of the regulation are quite complex, allowing for various types of exemptions to the AFV mandate. The primary target of the Act is government fleets of 50 or more vehicles, but some private fleets are also covered. Nonexempt fleets are typically required to maintain 75% or more AFVs.

[6] By contrast, a 2005 national survey sponsored by ethanol producer VeraSun Energy found that 68% of FFV owners were unaware of the E85 fueling capability (Reid 2006). The difference in these results strongly suggests that the concentration of E85 retail outlets is a primary factor driving awareness of FFV E85 compatibility. To address potential awareness concerns, since 2006 FFV manufacturers have included labels inside the fuel compartment to remind owners of the E85 capability.

[7] Historical fuel prices indicate that E85 is sometimes more economical on a \$/mile basis, but that more frequently FFV owners pay a premium for E85, a finding corroborated by Anderson (2011) and Petrolia et al. (2010). These results suggest that environmental and social benefits may be stronger drivers of FFV adoption than potential cost savings.

by state-dependent excise tax rates and the wholesale prices of ethanol and gasoline, which incorporate delivery transportation costs. Sourcing ethanol directly from a local refinery is thus advantageous for fuel retailers as it typically secures lower ethanol input prices. Fuel use will be determined by commuting patterns, which may vary by market and driver type (i.e., consumer or fleet). Similarly, willingness to pay for E85 is expected to vary by demographic factors and driver type. Drivers of fleet vehicles are presumably less price sensitive than consumers because fleet vehicles are subject to alternative fuel use mandates and these drivers are typically reimbursed for fuel expenses. Empirical specifications of the E85 supply model developed in §4.2 attempt to capture each of these determinants of profit contribution.

Fuel retailer decisions to enter the market for E85 must also consider required investments in E85 dispensing infrastructure. As with FFVs, gas stations must dispense E85 from storage tanks that are resistant to alcohol-related corrosion. Specialty hoses, nozzles, and handles are also required. Installation costs vary by configuration and geography, but a report by the National Renewable Energy Laboratory (NREL) estimates the median cost of a new storage tank to be \$59,153, and \$11,237 when converting an existing tank (National Renewable Energy Laboratory 2008). Several incentive programs are available to offset these installation costs. At the federal level, the Energy Policy Act of 2005 provides tax credits for up to 30% of the costs to install E85 infrastructure (capped at \$30,000). The four states studied here also offer E85 infrastructure incentives at various levels.[8] Infrastructure expenditures net of these subsidies are the retailer's (nonrecurring) fixed cost to enter the E85 market. Analogous to the FFV acquisition decision, the relatively small scale of net infrastructure costs, coupled with the backward compatibility of E85 dispensing equipment with gasoline, implies that dynamic considerations are less important in this setting than in those traditionally studied in the market entry literature. That is, because E85 dispensing infrastructure may be almost costlessly retrofitted to dispense gasoline, exiting the E85 market typically involves reallocating the E85 tank and pump to a traditional gasoline blend rather than discontinuing their use altogether. In this sense, infrastructure investments are sunk only to the extent they exceed the scrap value of a profit stream generated by using the tank/pump to dispense gasoline. In light of these considerations, I simplify the treatment of firms' E85 market entry and exit decisions by representing them as outcomes of a repeated static game. I return to this point during my concluding remarks in §8.

3. Data

Data for the study are a panel of nine yearly observations of 3,857 zip codes from Illinois, Indiana, Iowa, and Minnesota.[9] The data span the years 2001 to 2009 and are summarized in Table 1. I assemble the final data set by combining numerous data sources.

Vehicle data comes from R.L. Polk & Company, which compiles vehicle registration information from state departments of motor vehicles. The records provided are transaction level, and represent the *complete* population of FFV registrations for the period of study; these data are therefore one of the most comprehensive sets in the literature. Record attributes include the vehicle's zip code of registration, date of registration, and registrant type (e.g., consumer, government, or business). For the purpose of this study, I aggregate government and business FFV registrations into a category termed "fleet registrations." Government vehicle registrations comprise approximately 25% of the fleet category, with the balance coming from companies, excluding automobile dealerships and car rental companies.[10] From these data, I construct type-specific measures of FFV demand: the "flow" variables Q_1 and Q_2 capture per-period demand of consumer and fleet FFVs, and the "stock" variables *con_ffv_ib* and *flt_ffv_ib* capture cumulative demand (the FFV installed base).[11]

Data on E85 availability comes from the National Renewable Energy Laboratory (NREL). These data include the station name, station type (retail or private access), address, E85 introduction date, and station closure date (if applicable). However, no information is available on whether the station installed

[8] Minnesota: Grant amounts are discretionary but impose no cost sharing percentages or caps. Iowa: 50% cost sharing, capped at \$30,000. Illinois: 30% new, 50% conversion cost sharing, capped at \$30,000. Indiana: Grants of up to \$5,000 are available. Programs are subject to overall spending limits. See www.afdc.energy.gov/fuels/laws/3252 for details.

[9] Henceforth, "zip code" is taken to mean a Census 2000 Zip Code Tabulation Area (ZCTA). Zip codes are defined as a set of postal delivery addresses, whereas ZCTAs define unique spatial regions that closely correspond to the bounding polygons for zip codes.

[10] I exclude dealership registrations because these vehicles are inoperative inventory and thus do not reflect end-user vehicle demand or contribute to E85 consumption. Rental companies are closely aligned with auto manufacturers (e.g., Hertz is a wholly-owned subsidiary of Ford) and thus contractual obligations likely determine rental car demand. Moreover, rental companies generally operate their own fueling infrastructure and are thus unlikely to systematically contribute to local E85 consumption.

[11] Computation of the installed base entails an initial conditions problem, as my observations begin in 2001 whereas FFV production began in 1998. In addition, as the installed base intends to reflect FFVs in use, I must impute unobserved vehicle retirement. See §4.2 and Web Appendix D for more on this issue (available as supplemental material at http://dx.doi.org/10.1287/mksc.2014.0881).

Table 1 **Summary Statistics**

Variable	Name	Obs	Mean	Std. dev.	Min	Max
FFV registrations						
Consumer	Q_1	34,713	14.06	23.38	0	300
Fleet	Q_2	34,713	1.74	3.62	0	101
FFV installed base						
Consumer	*con_ffv_ib*	34,713	78.34	138.04	0	1,707
Fleet	*flt_ffv_ib*	34,713	9.14	18.95	0	454
E85 stations						
Retail	N_1	34,713	0.10	0.37	0	5
Private	N_2	34,713	0.00	0.07	0	1
Market characteristics						
Population (greater than 16 years)	P_1	34,713	5,298.62	9,076.82	6	80,216
Persons employed	P_2	34,713	3,032.11	6,823.27	1	127,465
Car dealerships	*Dealerships*	34,713	1.14	2.49	0	29
Gasoline stations	*Gas_stations*	34,713	1.92	4.14	0	33
Ethanol plants	*Refineries*	34,713	0.02	0.14	0	3
Rural market indicator	*Rural*	3,857	0.68	0.47	0	1
Median household income ('000)	*Income*	3,857	41.58	13.29	0	200
Median commute time (min)	*Commute*	3,857	24.46	6.22	0	61.6
Interstate highways	*Interstates*	3,857	0.2	0.45	0	3

new or converted existing infrastructure to enter the E85 market. I summarize E85 availability as the number of retail (N_1) and private access (N_2) fueling facilities in a (zip code/year) market. Finally, I assemble zip code statistics on consumer demographics (from the U.S. Census), businesses in the vehicle and fuel supply chains (from annual Census Zip Business Patterns), and highways (USGS shapefiles). These variables include measures of the potential market for FFVs among consumers (driving age population, P_1) and fleets (persons employed, P_2), factors influencing vehicle and fuel preferences (median income, rural/urban designation, commute time), availability of substitutes for FFVs and E85 (counts of auto dealerships and gas stations), key shifters of E85 variable costs (ethanol refineries), and a control for E85 demand from nonlocal FFVs (interstates).

In total, I observe 488,115 consumer FFV registrations, 60,437 fleet FFV registrations, and 840 E85 entry events, illustrating the relative richness of FFV demand data compared to the E85 supply data. To tease out the variation in E85 availability further, in Table 2 I tabulate observations of the number of E85 retailers per (zip code/year) market and compute the empirical probabilities of transitions in the number of E85 retailers. The table demonstrates, for example, that I have 2,415 observations of markets with a single E85 retailer, 367 observations of markets with two E85 retailers, and so on. Given that there was a single E85 retailer active in the market in the previous period ($t-1$), there is a 1% probability of observing the firm exit the market in the current period (t), a 94% probability the firm remains the sole incumbent, a 5% probability a second firm enters the market, and so on.[12] The data clearly indicate that market configurations tend to be highly persistent over time (entry and exit are rare events) and that a local monopoly is the most common structure in markets where E85 is present.[13] In the last two rows of the table I also report the average size of the consumer and fleet installed bases of FFVs, conditional on the number of E85 retailers. Note that the installed bases increase as N_1 increases. While this relationship is expected in the presence of a network effect, it is merely suggestive since these averages combine cross-sectional and time-series variations and do not control for factors such as market population.[14]

Before developing the model, I present model-free evidence for the existence of a network effect. The expected patterns are that the number of E85 retailers (N_1) is positively correlated with the installed base of consumer and fleet FFVs and that flex-fuel demand

[12] In the table, I adjust the prior period number of E85 retailers such that it is net of station closures. Thus, exit events correspond to "product exit", i.e., reallocating E85 pumps and tanks to gasoline. The two types of exit occur with nearly equal frequency in the data: I observe 33 "product exits" and 34 "station closures." See §4.2 for more discussion of this issue.

[13] In a slight abuse of terminology, a local monopoly in the current context refers to having no other direct E85 competitors in the same zip code. In general, there may be other substitute fuel (gasoline) retailers in the market. To give a sense of the geographic scale of what is meant by "local," a zip code has on average an area equivalent to a circle of a five mile radius.

[14] A similar pattern of positive correlation between the FFV installed base and the number of E85 retailers is evident in plots of the spatial distribution of FFVs and E85. I provide a time sequence of these maps in Web Appendix A.

Table 2 Retail E85 Stations per Market (N_{1mt}) and Empirical Markov Transitions ($N_{1m,t-1} \to N_{1mt}$)

	N_{1mt}						
$N_{1m,t-1}$	0	1	2	3	4	5	Obs
0	0.98	0.02	0.00	0.00	0.00	0.00	32,428
1	0.01	0.94	0.05	0.00	0.00	0.00	1,939
2	0.00	0.02	0.87	0.10	0.01	0.00	268
3	0.00	0.00	0.02	0.86	0.12	0.00	58
4	0.00	0.00	0.00	0.00	0.89	0.11	19
5	0.00	0.00	0.00	0.00	0.00	1.00	1
Obs	31,812	2,415	367	87	29	3	34,713
Consumer FFV installed base	64.7	194.8	369.4	411.5	574.7	917.7	
Fleet FFV installed base	7.2	26.3	48.8	60.7	88.1	118.3	

(Q_1, Q_2) is positively correlated with the number of E85 retailers (N_1). In Table 3 I report descriptive linear regressions of the dependent variables Q_1, Q_2, and N_1 as a function of the relevant endogenous variables. The first column for each dependent variable reports the regression results with no controls or instruments. The second and third columns add fixed effects to control for unobervables potentially correlated with the endogenous regressors of interest. Zip code fixed effects control for unobserved market characteristics such as consumer environmental preferences and regional E85/FFV incentives or awareness campaigns. Zip code time trends control for factors such as population growth and gentrification. Year effects control for variation in the variety of FFV models available, average fuel prices and input costs, and changes in federal E85/FFV incentives. Similarly, state/year fixed effects control for changes to state-level E85/FFV incentives and mandates as well as regional variation in E85/FFV awareness. In the third column for each dependent variable, an instrumental variable regression is reported, where the instruments used are two period lags of the endogenous variables. The identification assumption is that the econometric error terms are serially uncorrelated. I test the validity of this identification assumption (in the context of the full model) in §6; my purpose here is simply to demonstrate the robustness of descriptive regression estimates to alternative identification assumptions. The fact that all coefficients on the FFV installed bases and N_1 have the expected (positive) signs and are statistically significant (with the exception of instrumented fleet demand) provides strong prima facie evidence for the presence of a network effect.

4. Model

The model considers the decisions of three types of economic agents: fuel retailers, consumers, and fleets. The conceptual framework is a simultaneous move game, played each period by agents in the local market who base their decisions on complete knowledge of current market conditions only partially observed by the econometrician. The moves taken by agents in the game are as follows: (1) consumers/fleets choose whether to purchase an FFV, (2) fuel retailers decide whether to enter the E85 market, and (3) FFV owners set their level of E85 consumption and entering retailers compete in quantities to set output levels. An output setting game is chosen as competition in prices among identical firms' results in the familiar Bertrand paradox where all firms price at marginal cost. Firm competition is strategic in that each firm's entry decision influences the level of profits realized by all firms in the market. As the flex-fuel purchase decisions of consumers and fleets influence firm profits by increasing the size of the market for E85, the market equilibrium will be defined over the action space of all agents, not just firms. The solution concept of the game is a symmetric Nash equilibrium in pure strategies.

Throughout the model development, I use the following notational conventions: I index markets (zip codes) by m, wider regions (states) by r, individuals by i, and time periods (years) by t. I index equations by $k \in \{1, 2, 3\}$ for consumer flex-fuel demand, fleet flex-fuel demand, and E85 market entry, respectively.

4.1. Flex-Fuel Demand

4.1.1. Consumers. The flex-fuel demand system measures market-specific feedback in FFV adoption from E85 entry in a robust but parsimonious fashion. The model must draw inference from counts of FFV registrations that do not convey transaction prices or current vehicle holdings, among other important factors. I must therefore abstract from numerous details that would be pertinent in a more focused analysis of automotive demand. With these caveats in mind, I model consumer demand for FFVs as a binary choice. In every period, consumers choose between purchasing an FFV (1) and the outside alternative (0), which

Table 3 Descriptive Regressions of Model Dependent Variables

Dependent variable	Q_1	Q_1	Q_1	Q_2	Q_2	Q_2	N_1	N_1	N_1
Retail E85 stations (N_1)	21.17	3.06	3.47	3.32	0.36	0.29			
	(0.32)	(0.67)	(1.51)	(0.05)	(0.15)	(0.79)			
Private E85 stations (N_2)				5.45	−0.41	0.77			
				(0.26)	(0.71)	(5.36)			
Consumer FFV installed base							0.0004	0.0009	0.0011
							(0.0000)	(0.0001)	(0.0002)
Fleet FFV installed base							0.0051	0.0032	0.0037
							(0.0002)	(0.0004)	(0.0013)
Observations	34,713	30,856	26,999	34,713	30,856	26,999	34,713	30,856	26,999
Zip FE	N	Y	Y	N	Y	Y	N	Y	Y
Year FE	N	Y	Y	N	Y	Y	N	Y	Y
State/year FE	N	Y	Y	N	Y	Y	N	Y	Y
Zip time trends	N	Y	Y	N	Y	Y	N	Y	Y
Instruments	N	N	Y	N	N	Y	N	N	Y

is normalized to have zero utility in expectation. That is, consumers are assumed to be in the market for an FFV every period, allowing for the possibility of replacement sales.[15] The choice-specific utilities for an individual consumer are assumed to take the following form:

$$U^1_{imt} = \alpha_1 N_{1mt} + \beta_1' X_{1mt} + \delta_{1m} + \omega_{1t} + \zeta_{1rt} + t\nu_{1m} + \varepsilon_{1mt} + \eta^1_{imt} \equiv \bar{U}_{1mt} + \varepsilon_{1mt} + \eta^1_{imt} \quad (1a)$$

$$U^0_{imt} = \eta^0_{imt}. \quad (1b)$$

The utility of choice 1 appearing in Equation (1a) captures the consumer's valuation of an FFV as a function of the number of E85 retailers operating in the market, N_{1mt}. The associated α_1 parameter is the primary coefficient of interest, which captures feedback in E85 availability on FFV adoption. The β_1 parameters capture the influence of exogenous observables (X_{1mt}) such as demographic characteristics and the number of auto dealerships in the market. As in the descriptive regressions of the previous section, the utility specification (1a) uses strong controls for unobservables through the inclusion of: (i) market fixed effects (δ_{1m}), (ii) year fixed effects (ω_{1t}), (iii) state/year fixed effects (ζ_{1rt}), and (iv) zip code specific time trends ($t\nu_{1m}$). ε_{1mt} is a market and period specific shock to FFV utility common to all consumers. The terms η^0_{imt} and η^1_{imt} capture individual *i's* idiosyncratic preferences for the choice options, which are assumed to be independent and identically distributed extreme value shocks. The extreme value assumption implies that choice-specific market shares follow the standard logit formulas:

$$H^1_{1mt} = \frac{Q_{1mt}}{P_{1mt}} = \Pr[U^1_{imt} > U^0_{imt}] = \frac{\exp(\bar{U}_{1mt} + \varepsilon_{1mt})}{1 + \exp(\bar{U}_{1mt} + \varepsilon_{1mt})} \quad (2a)$$

$$H^0_{1mt} = \frac{P_{1m} - Q_{1mt}}{P_{1mt}} = \Pr[U^1_{imt} \le U^0_{imt}] = \frac{1}{1 + \exp(\bar{U}_{1mt} + \varepsilon_{1mt})} \quad (2b)$$

where H^1_1 and H^0_1 are the choice shares of FFVs and the outside option, Q_1 is the current period FFV registrations, and P_1 is the driving age population of the market. As shown by Berry (1994), the market share equations above may be transformed into the following linear expression for the log-odds of consumer FFV purchase, H_1:[16]

$$H_{1mt} = \ln(H^1_{1mt}) - \ln(H^0_{1mt}) = \ln \frac{Q_{1mt}}{P_{1mt} - Q_{1mt}} = \bar{U}_{1mt} + \varepsilon_{1mt}. \quad (3)$$

4.1.2. Fleets. As discussed in §2, fleet operators face different incentives than consumers when considering vehicle purchases; thus I model the populations separately. Modeling fleet vehicle demand

[15] I allow replacement sales to be consistent with my computation of the installed base of FFVs, which explicitly accounts for vehicle retirement. In practice, this assumption has little effect on demand estimates, as market populations are large in comparison to cumulative flex-fuel sales (in my sample, the FFV penetration rate averages about 2% of the market population).

[16] In empirical specifications, I calculate H_{1mt} using $H_{1mt} = \ln((\kappa + Q_{1mt})/(\kappa + P_{1mt} - Q_{1mt}))$ as this avoids the technical problem of a negative infinite log-odds ratio when $Q_{1mt} = 0$ (which occurs in 15% of my observations). For the correction I set $\kappa = 0.5$, as this value is shown by Pettigrew et al. (1986) to be the bias-minimizing value. Using the asymptotic value derived by Pettigrew et al. (1986), I estimate the bias in my application to be $\sim -10^{-5}$, which is negligible in comparison to the mean of the log-odds ratio of −5.7.

involves additional challenges due to the complexity of the underlying acquisition process and limitations of the data. Fleet acquisitions are often centrally coordinated and therefore demand is characterized by infrequent purchases of multiple vehicles. Depending on where vehicles are registered (e.g., organization headquarters or employee home) and the extent to which employees live outside their employer's market, the fleet demand process may also span multiple markets. Fully accounting for these features would require detailed data on organizations operating vehicle fleets and the ability to associate observed fleet vehicle registrations with a particular organization. Unfortunately these data are not available, necessitating a more abstract treatment of fleet FFV demand. My empirical strategy is therefore to choose a convenient fleet demand specification and to demonstrate robustness of the measured network effects to alternative specifications. I explore alternative fleet demand specifications in Web Appendix E. The measurement of network effects is substantially robust to the choice of specification.

Accordingly, I preserve symmetry with the consumer FFV demand specification and use an aggregate logit specification identical in form to Equation (3)

$$\begin{aligned} H_{2mt} &= \ln \frac{Q_{2mt}}{P_{2mt} - Q_{2mt}} = \bar{U}_{2mt} + \varepsilon_{2mt} \\ &= \alpha_{21} N_{1mt} + \alpha_{22} N_{2mt} + \beta_2' X_{2mt} + \delta_{2m} + \omega_{2t} + \zeta_{2rt} \\ &\quad + t\nu_{2m} + \varepsilon_{2mt}. \end{aligned} \tag{4}$$

This specification effectively treats fleets as a collection of individuals who make independent FFV adoption decisions. In light of the abstraction from bulk-buying behavior, parameter estimates from this model should not be interpreted as primitives of fleet utility. In (4), I take the number of persons *employed* (P_2) as the measure of market size. This definition seems plausible, as fleet vehicles are typically assigned to employees and markets with higher employment should, all else equal, contain a larger number of fleet vehicles. Market fixed effects and time trends are particularly important in this context, as they control for unobserved heterogeneity in the true fleet FFV market potential. Equation (4) differs from (3) in one other respect: the number of private-access E85 refueling stations (N_2) also enters the specification. N_2 is potentially endogenous since there may be common unobservables (e.g., regional incentive programs) driving the construction of these dedicated E85 facilities and fleet FFV demand. I use instrumental variables to address the potential endogeneity of N_2 during estimation.

4.2. E85 Market Entry

Fuel retailer entry into the E85 market is modeled in the tradition of the seminal works by Bresnahan and Reiss (1990, 1991, 1994). The unifying theme of these papers is that the number of firms in a market, N, is modeled as the outcome of a two-stage game of complete information played among a set of E firms with ex ante symmetric profit functions. In the first stage, firms make strategic decisions to enter (or continue participating in) the market, anticipating the ensuing competition (in quantities or prices) in the second stage. Recognizing that firm profits decline as competition increases, observations of N reflect bounds on the latent profit function, which may be recovered through ordered categorial variable regression techniques.

My specification follows Bresnahan and Reiss (1994) in its application to geographic panel data. The model distinguishes between potentially entering firms and incumbent firms in that only newly entering firms incur nonrecurring entry costs. Potential entrants enter the market when their expected discounted profits exceed entry costs. Incumbents continue to operate as long as expected discounted profits exceed the scrap value obtained from reallocating E85 tanks and pumps to dispensing gasoline. To formalize the framework in the context of my application, let Π^N represent the expected discounted profit contribution (i.e., Π^N excludes fixed costs) of selling E85 for each retailer in a market where N firms are operating. Let F represent the nonrecurring cost of entry, i.e., the fixed cost of E85 infrastructure investments. The marginal (Nth) potential entrant will therefore enter the market when $\Pi^N \geq F$. I normalize the scrap value of exiting the market to zero, so that the marginal incumbent will exit the market when its expected discounted profits are less than zero ($\Pi^N < 0$), and continue operating as long as expected profits are non-negative ($\Pi^N \geq 0$).[17] Finally, let $\tilde{N}_{1mt}$ denote the *adjusted* number of E85 retailers in a market, where $\tilde{N}_{1mt} = N_{1mt} - (\#\ of\ E85\ station\ closures\ in\ t)$. I define this adjusted quantity to distinguish between voluntary exit from the E85 product market and involuntary exit that occurs when a fuel station shuts down entirely.[18] Given that E85 comprises a small portion

[17] Nonrecurring entry costs, scrap values, and the level of unobserved firm profits cannot be separately identified. Thus, normalizing one of these quantities is required. Normalizing the scrap value to zero does not affect the measurement of entry thresholds (the number of FFVs required to support a given number of E85 retailers), which is the key object of inference in the entry model.

[18] A related issue is distinguishing between an existing gas station adding an E85 tank/pump and an entirely new gas station entering the market with an E85-compliant tank/pump. The NREL data do not provide these details, but using secondary sources I obtained opening dates for 507 stations that carry E85. Of those, only 22 offered E85 in the first year of establishment. This suggests most E85 entry comes from existing gas stations.

(typically $< 5\%$) of station revenues, I assume station closures are exogenous events essentially unrelated to E85 product decisions. Under this definition, the quantity $N_{1mt} - \tilde{N}_{1m,t-1}$ reflects the number of firms that voluntarily incur or forfeit E85 infrastructure investments in period t, which is the relevant measure of entry/exit linked to the E85-specific profit function. An observation of N firms in the data can therefore correspond to one of three conditions:

1. Net entry: $N_{1mt} > \tilde{N}_{1m,t-1}$. At least one potential entrant has entered the market. This condition implies that per-firm expected discounted profits were sufficient to support the marginal (Nth) entrant, but insufficient to support an additional ($(N+1)$th) entrant. That is, $\Pi^N_{mt} \geq F_{mt}$ and $\Pi^{N+1}_{mt} < F_{mt}$.
2. No change: $N_{1mt} = \tilde{N}_{1m,t-1}$. The number of firms stayed the same. Thus, per-firm expected discounted profits were sufficient to support the N incumbents, but insufficient to support an additional ($(N+1)$th) entrant. That is, $\Pi^N_{mt} \geq 0$ and $\Pi^{N+1}_{mt} < F_{mt}$.
3. Net exit: $N_{1mt} < \tilde{N}_{1m,t-1}$. At least one incumbent has exited the market. Per-firm expected discounted profits were sufficient to support N incumbents, but insufficient to support $N+1$ incumbents. That is, $\Pi^N_{mt} \geq 0$ and $\Pi^{N+1}_{mt} < 0$.

The above conditions may be compactly expressed as $\Pi^N_{mt} \geq I(N_{1mt} > \tilde{N}_{1m,t-1})F_{mt}$ and $\Pi^{N+1}_{mt} < I(N_{1mt} \geq \tilde{N}_{1m,t-1})F_{mt}$, where the indicator function $I(\cdot)$ evaluates to 1 if the enclosed statement is true. These inequalities, combined with distributional assumptions on unobservables, are used to recover parameters characterizing firm profits, which can in turn be used to quantify the influence of indirect network effects.

I assume that expected discounted profits are multiplicatively separable in market size (S) and per-capita variable profits (V) and can be given by the following reduced form:

$$\Pi^N_{mt} = S_{mt}\frac{V_{mt}}{(1+N_{1mt})^2} + \varepsilon_{3mt}. \tag{5}$$

This specification is motivated by assumptions of linear per-capita demand for E85, constant variable costs, and second stage Cournot competition in quantities. I derive this reduced form profit function for a one-period game in Web Appendix B. In (5), the market size S is measured in FFV units, and per-capita variable profits V capture the expected discounted profit stream from the vehicles in S. Following Bresnahan and Reiss (1991), I assume that unobservable firm profits enter Π^N linearly as a common shock denoted by ε_{3mt}.

To complete the econometric specification, I must clarify how S, V, and F are parameterized. I represent nonrecurring fixed costs as a linear function of exogenous market-level cost shifters (X_3): $F_{mt} = \phi' X_{3mt}$. I parameterize per-capita variable profits as $V_{mt} = \exp(\psi' X_{4mt})$, where X_4 is a vector of exogenous E85 variable cost and demand shifters. Forcing the positivity of variable profits through exponentiation is consistent with theory and improves convergence of the estimation algorithm. Finally, I assume the market size S may be represented as a linear function of the installed base of consumer FFVs, the installed base of fleet FFVs, and other exogenous shifters (X_5)

$$S_{mt}(N_{1mt}) = \left(\sum_{\tau=-\infty}^{t} w_{t-\tau} Q_{1m\tau}\right) + \gamma\left(\sum_{\tau=-\infty}^{t} w_{t-\tau} Q_{2m\tau}\right) + \lambda' X_{5mt}. \tag{6}$$

In (6) above, the installed base terms are computed as a weighted sum of cumulative FFV demand in a market. The weight w_k represents the survival rate for a vehicle that is k years old (where $w_0 \equiv 1$), so that the weighted sum approximates the number of on-road FFVs after accounting for vehicle retirement.[19] In addition to observed FFV demand from 2001–2009, the summation includes imputed FFV demand for 1998–2000. Details of the imputation procedure may be found in Web Appendix D. The coefficient on the installed base of consumer FFVs is normalized to one for identification purposes, as the (fully parameterized) expression for variable profits interacts with the market size in Equation (5). This normalization converts E85 demand into units of consumer FFVs and thus implies that the γ coefficient in Equation (6) may be interpreted as the estimated number of fleet FFVs required to match the profit contribution of one consumer FFV. I write S as a function of N_1 to emphasize that the market size is an implicit function of N_1 through current period FFV demand, Q_1 and Q_2. This relationship is the source of simultaneity in the model. By extension, (6) makes clear that feedback realized in the current period will persist into future periods through the weighted summation of the installed base.

Putting the aforementioned specifications together, the model for N_1 taken to data is

$$N_{1mt} = \sum_{k=0}^{E} k I(\Pi^k_{mt} \geq I(N_{1mt} > \tilde{N}_{1m,t-1})F_{mt}) \cdot I(\Pi^{k+1}_{mt} < I(N_{1mt} \geq \tilde{N}_{1m,t-1})F_{mt}) \tag{7a}$$

$$\Pi^k_{mt} = \begin{cases} 0 & k = 0 \\ S_{mt}(k)\dfrac{V_{mt}}{(1+k)^2} + \varepsilon_{3mt} & 0 < k \leq E \\ -\infty & k = E+1. \end{cases} \tag{7b}$$

[19] Weights are assumed to be time stationary and as quoted in the Department of Transportation's Transportation Energy Data Book, Edition 25, Table 3.9 (Davis and Diegel 2006). As the weights decline very slowly in duration, measurement of the market entry thresholds in Table 6 is not very sensitive to their use. Omitting them altogether increases entry thresholds by less than 4%.

4.3. Equilibrium

A challenge to inference in empirical models of strategic interactions, such as the market entry game considered here, is the presence of multiple equilibria. If a model admits multiple equilibria, the mapping from parameters to game outcomes is not necessarily unique, complicating both estimation and counterfactual evaluation. Bresnahan and Reiss (1990) demonstrate how modeling the number (rather than identities) of entering firms resolves multiplicity when firms are identical and profits strictly decrease in the number of entrants (due to competition). In my application, entry can expand the market and thus potentially increase profits, admitting the possibility of a nonmonotonic relationship between profits and the number of firms in the market. I ensure a unique equilibrium by ruling out *net* profit increases from entry by imposing the constraint that $\partial \Pi^N/\partial N_1 \leq 0$ hold for all markets when estimating Equation (7). That is, increases in firm profits from an expanded market size (via the network effect in $S(N_1)$) are not allowed to be larger than decreases to profits from additional competition (via the $1/(1+N_1)^2$ term entering variable profits). The constraint, which is derived in Web Appendix C, takes the following form:[20]

$$G(N_1) \equiv \frac{\alpha_1 Q_1}{1+\exp(H_1)} + \frac{\gamma \alpha_{21} Q_2}{1+\exp(H_2)} - \frac{2S}{1+N_1} \leq 0. \quad (8)$$

Formally then, the Nash equilibrium for the model is defined by the number of operating E85 retailers N_1^*, the number of consumer FFV sales Q_1^*, and the number of fleet FFV sales Q_2^* that simultaneously satisfy Equations (3), (4), (7), and (8).

5. Estimation

As explained below, I use a simulation-based technique to estimate the E85 entry model, which in turn requires a distributional assumption on the econometric error terms. Building on the standard assumption of normally distributed shocks to firm profitability in the entry literature, I accommodate correlated shocks between E85 profits (ε_3) and FFV demand ($\varepsilon_1, \varepsilon_2$) by assuming that ε is distributed multivariate normal.[21] The variance of the entry model error term ε_3 is normalized to one to identify the unobserved scale of firm profits. The system to be estimated thus has the following structure:

$$\begin{aligned} &Q_1(X, N_1, \varepsilon_1; \theta_1), Q_2(X, N_1, N_2, \varepsilon_2; \theta_2), \\ &\quad N_1(X, Q_1, Q_2, \varepsilon_3; \theta_3), (\varepsilon_1, \varepsilon_2, \varepsilon_3) \sim N(0, \Sigma), \\ &\quad \Sigma_{33} \equiv \sigma_3 = 1, \end{aligned} \quad (9)$$

where I collect all exogenous covariates into X and equation-specific parameters into the θ terms. I devise a two-step estimation procedure to consistently estimate the system (9). As the system is estimated sequentially, I compute standard errors using a panel bootstrap procedure to account for measurement error in the first step (FFV demand) parameters when estimating E85 entry model in the second step. The two-step estimation procedure is performed on 30 samples where market histories are drawn with replacement; standard errors are given by the standard deviations of the parameter estimates.

In the first step of the procedure, I leverage methods from the panel data literature to estimate the flex-fuel demand parameters (θ_1 and θ_2) using a rich fixed effects specification to control for unobservables and instruments to address simultaneity concerns. Estimation proceeds by working with the expressions for the log-odds of FFV purchase, Equations (3) and (4), which are linear in parameters and related to the first two expressions in (9) above via $Q_1 = P_1(\exp(H_1)/(1+\exp(H_1)))$ and $Q_2 = P_2(\exp(H_2)/(1+\exp(H_2)))$. I recover the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ by applying a variant of the System GMM estimator of Blundell and Bond (1998). As Blundell and Bond (1998) note, System GMM is well suited for the estimation of panel data models where: (i) fixed effects for the panel variable enter as controls, (ii) supplemental exogenous shifters of the endogenous regressors are unavailable, and (iii) the evolution of endogenous regressors approximates a random walk. These are precisely the conditions I face in the current application.[22] SGMM operates under a sequential exogeneity assumption, i.e., that current period econometric error terms are uncorrelated with prior realizations of the endogenous variables but are potentially correlated with current and future values of those variables. This assumption in turn implies that lagged values of an endogenous variable may be used as instruments for that variable in the current period. Two tests are available to substantiate the validity of

[20] The constraint does not bind at the estimated parameter vector, which is unsurprising given the modest rate of E85 diffusion.

[21] As will be seen, the normality assumption is not necessary for or used during FFV demand estimation, which uses GMM. Ex post, the assumption seems reasonable as the estimated FFV demand residuals closely approximate a normal distribution.

[22] With market fixed effects in the model specification, supplemental instruments must vary in a robust fashion over time within the same zip code. Theoretically appealing candidates, such as shifters of E85 infrastructure costs (which should be correlated with E85 entry but not FFV demand), exhibit little or no temporal variation. As Table 2 makes clear, evolution of the number of E85 retailers is well characterized by a first order Markov process.

the System GMM estimator. First, as is standard with GMM, the Hansen J (Hansen 1982) test is available to determine whether the lagged endogenous variables are properly excluded from the estimation equation. Second, residuals from the regression may be tested for serial autocorrelation, which, if present, would violate the sequential exogeneity assumption. I report both tests in the results of §6.

Implementation of SGMM requires forming two sets of moment conditions, one involving the estimation equation in first-differenced form (the "differences" equation) and one involving the undifferenced form (the "levels" equation). The SGMM estimator uses both sets of conditions to improve econometric efficiency. In the differences equation, regressors are transformed (differenced) to purge the fixed effects. Levels of endogenous variables lagged by two or more periods are available as instruments for those variables in differenced form. In the levels equation, instruments are transformed (differenced) to make them orthogonal to the fixed effects. Differences of endogenous variables lagged by one or more periods are then available as instruments for the variables in level form. Translating these statements into an estimator for the consumer FFV demand equation begins by defining instrument sets for the difference and level equations as $Z_{1mt} \equiv (\Delta X_{1mt}\ N_{1m,\,t-2}, \ldots, N_{1m1})$ and $\Delta Z_{1mt} \equiv (X_{1mt}\ \Delta N_{1m,\,t-1}, \ldots, \Delta N_{1m2})$ where Δ is the first-difference operator. In a slight abuse of notation, I let X_{1mt} incorporate indicator variables representing year fixed effects, state/year fixed effects, and zip code time trends in addition to the observable regressors described in Equation (3). The moment conditions for the GMM objective are then

$$\mathbb{M}_{1mt} = \begin{pmatrix} Z_{1mt}(\Delta\varepsilon_{1mt}) \\ (\Delta Z_{1mt})\varepsilon_{1mt} \end{pmatrix}$$

and the GMM estimator $\hat{\theta}_1 = \arg\min_{\theta_1} \sum_{m,\,t} \mathbb{M}'_{1mt}\mathbb{W}_1 \cdot \mathbb{M}_{1mt}$ is formed as a sample analog of the identification assumption $E[\mathbb{M}_1] = 0$. The optimal weight matrix $\mathbb{W}_1$ is inversely proportional to the variance of the moments and estimated via the usual two-step procedure assuming independent moments in the first step. The fleet FFV demand parameters are estimated analogously by $\hat{\theta}_2 = \arg\min_{\theta_2} \sum_{m,\,t} \mathbb{M}'_{2mt}\mathbb{W}_2\mathbb{M}_{2mt}$, where $\mathbb{M}_{2mt}$ is constructed similarly to $\mathbb{M}_{1mt}$ but also includes in Z_{2mt} (ΔZ_{2mt}) lags (lagged differences) of private E85 facilities (N_2) as instruments to accommodate the potential endogeneity of N_2.

In the second step of the overall procedure, I use simulation to estimate the E85 entry model. The key insight motivating the estimator is that conditional on FFV demand estimates $(\hat{\theta}_1, \hat{\theta}_2)$ and a draw of the structural error terms ε, the system of Equations (3), (4), and (7) may be solved simultaneously to generate the model-predicted number of E85 retailers in equilibrium (N_1^*) for any value of the entry model parameters (θ_3). Explicitly solving the reduced form fully accounts for the endogeneity of FFV demand during the second step estimation, as all FFV demand dependencies, including cross-equation correlation in the structural error terms, are consistently reflected in the model-predicted number of E85 retailers.

To promote numerical convergence of the second step estimation, a smooth simulator is generated by averaging predicted entry values over L draws of the error terms, i.e., $N^*_{1mt} \equiv (1/L)\sum_{l=1}^{L} N^*_{1mtl}$. I provide further details on this computation in §5.1 below. With a smooth simulator available, estimation of θ_3 proceeds by minimizing the difference (in the sum of squares sense) between the observed number of E85 retailers (N_{1mt}) and the model-predicted number of E85 retailers (N^*_{1mt}). One final wrinkle is accounting for simulation bias. It has long been recognized in the literature that "plugging in" an unbiased simulator to a nonlinear objective function results in a biased estimator for a fixed number of simulation draws. To achieve an unbiased estimator, the analyst must either apply a bias correction or use a "large" number of simulations. I leverage the result of Laffont et al. (1995), who derive the asymptotic distribution of the bias-correcting term for the least squares objective and show that it is given by $(1/L)E[\mathrm{Var}(X)]$, where X is the simulator. Thus, my simulated nonlinear least squares estimator of θ_3 is given by[23]

$$\hat{\theta}_3 = \underset{\theta_3:\,G(N^*_{1mt})\le 0}{\arg\min} \sum_{m,\,t}\Bigg[(N_{1mt} - N^*_{1mt})^2 - \frac{1}{L(L-1)} \cdot \sum_{l=1}^{L}(N^*_{1mtl} - N^*_{1mt})^2\Bigg]. \tag{10}$$

5.1. Computation of Market Entry Simulator, N_1^*

The first step to composing the simulator is to obtain draws of the market shocks. To promote estimation convergence, a set of L independent draws from a standard trivariate normal distribution are taken once and stored at the beginning of the algorithm. Desired correlation patterns are induced on those draws in each optimization iteration by forming the product with the Cholesky root of the covariance matrix implied by θ_3. Note that because the covariance parameters of the FFV demand errors ($\sigma_1, \sigma_2, \rho_{12}$) can be estimated from the first step residuals (e.g., $\hat{\rho}_{12} =$

[23] In my empirical work, I take $L = 30$ draws for each of the 30 bootstrap replications. The constraint in Equation (10), which is expressed in terms of parameters and observables in Equation (8), is implemented using a penalty function during estimation.

$\text{Corr}(\hat{\varepsilon}_1, \hat{\varepsilon}_2)$), the cross-equation correlations ρ_{13} and ρ_{23} are the only elements of Σ that enter θ_3.[24]

With a set of properly correlated draws available, the next step is to compute firm profits with a *conjectured* number of E85 retailers in the market. The key insight here is that for the lth draw, the conjectured unit sales of consumer FFVs in a market with n E85 retailers may be written as $\hat{Q}_{1mtl}(n) = P_{1mt}(\exp(H_{1mt} - (\hat{\alpha}_1 N_{1mt} + \hat{\varepsilon}_{1mt}) + (\hat{\alpha}_1 n + \varepsilon_{1mtl})))/(1 + \exp(H_{1mt} - (\hat{\alpha}_1 N_{1mt} + \hat{\varepsilon}_{1mt}) + (\hat{\alpha}_1 n + \varepsilon_{1mtl})))$, and similarly for fleet unit sales, $\hat{Q}_{2mtl}(n)$. The operands of the exponential functions in these expressions are the observed log-odds of FFV purchase, with the effects of observed entry and market shocks replaced with their conjectured effects. This formulation serves two purposes. First, it ensures that the predicted values $\hat{Q}_1(n)$ and $\hat{Q}_2(n)$ incorporate the market fixed effects δ, which are not estimated explicitly. Second, replacing $\hat{\varepsilon}_{1mt}$ with ε_{1mtl} ensures that the econometric errors are consistent with the assumed data generating process for each conjectured value of n. That is, the process accounts for the fact that the values of correlated cross-equation shocks are endogenously determined along with the dependent variables (it would therefore be inappropriate to condition on realized shocks at alternative levels of E85 availability). To evaluate conjectured profits, $\hat{Q}_1(n)$ and $\hat{Q}_2(n)$ may be substituted into Equation (6) for market size and subsequently to (7b). The equilibrium number of firms is then determined by iteratively searching over the conjectured number of firms n (starting at $n = 0$) until the unique value that satisfies condition (7a), i.e., $(\Pi^n_{mtl} \geq F_{mt}(n > \tilde{N}_{1m,\,t-1})) \cap (\Pi^{n+1}_{mtl} < F_{mt}(n \geq \tilde{N}_{1m,\,t-1}))$, is found. During estimation, I take $E = \max[N_{1mt}] = 5$ to match the maximum observed number of entering E85 retailers.

5.2. Identification

The primary focus of my empirical strategy to identify the FFV demand parameters θ_1 and θ_2 is to absorb as many sources of confounding variation as possible through extremely rich nonparametric controls. With the inclusion of these controls, the key parameters of interest (α_1 and α_{21}) will be identified from within-market FFV share differences over time that (a) coincides with changes in E85 availability, and (b) are unrelated to market-specific trends, common year effects, or common state/year effects. Additional identification comes from the assumption that market shocks are orthogonal to lagged values of the number of E85 retailers operating in the market. While these identification assumptions seem reasonable given the institutional context, I cannot rule out all potential confounds. For example, higher-order (i.e., nonlinear) unobserved market-specific trends affecting FFV demand and E85 availability could be observationally confounded with a network effect. Similarly, an unobserved trend affecting one side of the market coupled with rational expectations by the other side could produce the appearance of a network effect. An example might be that vehicle manufacturers attempt to push FFVs on consumers who have no native preference for E85 compatibility and that fuel retailers respond by offering E85 in markets where they recognize this strategy being implemented.[25] If such systematic FFV distribution policies exist and result in nonlinear trends in FFV sales (I control for unobserved linear trends in FFV demand), they would lead to an inappropriately optimistic estimate of the network effect. While I view this particular scenario as rather unlikely given that most state franchise laws prohibit manufacturers from requiring dealerships to accept specific vehicle models, it illustrates the potential limitations of my empirical approach.[26]

Turning to the E85 entry model, identification of the parameters θ_3 rests on a combination of assumptions. As with all latent variable models, identification of θ_3 is driven in part by assumptions on the distribution of the econometric errors and the functional form of the latent construct. Exclusion restrictions provide an additional source of nonparametric identification. As the empirical specifications discussed in §6 make clear, in addition to several fixed effects, the observable regressor *dealers* enters the FFV demand specification but is omitted from the E85 entry model. The number of auto dealerships shifts FFV demand through variation in the availability of FFVs and substitutes for FFVs. As there is no reason to presume a connection between auto dealerships and E85 supply considerations, *dealers* are plausibly excluded from the E85 profit function. As to separate identification of infrastructure costs and variable profit factors, inspection of (7) suggests that infrastructure cost factors (ϕ) are identified by the mean values of N_1 (conditional on regressors X_3), while variable profit factors (ψ) will be identified through interactions of X_4 with the market size.

[24] That is, market shocks are generated by $\varepsilon(\rho_{13}, \rho_{23}) = C\kappa$ where $\kappa \sim I_3$ and

$$\Sigma = CC' = \begin{pmatrix} \hat{\sigma}_1^2 & \hat{\rho}_{12}\hat{\sigma}_1\hat{\sigma}_2 & \rho_{13}\hat{\sigma}_1 \\ \hat{\rho}_{12}\hat{\sigma}_1\hat{\sigma}_2 & \hat{\sigma}_2^2 & \rho_{23}\hat{\sigma}_2 \\ \rho_{13}\hat{\sigma}_1 & \rho_{23}\hat{\sigma}_2 & 1 \end{pmatrix}.$$

[25] I thank an anonymous reviewer for bringing this point to my attention.

[26] As an example, per Illinois code §815 ILCS 710/4, it is illegal for a vehicle manufacturer to coerce, or attempt to coerce, any motor vehicle dealer "to accept, buy or order any motor vehicle or vehicles, appliances, equipment, parts or accessories therefor, or any other commodity or commodities or service or services which such motor vehicle dealer has not voluntarily ordered or requested except items required by applicable local, state or federal law."

Table 4 Main Estimation Results

	FFV demand (H_1, H_2)		E85 Entry (N_1)		
	Consumer	Fleet	Variable profit	Infrastructure costs	Market size
Retail E85 stations (N_1)	0.059 (0.027)	0.064 (0.031)			
Private E85 stations (N_2)		0.026 (0.244)			
Rural	0.022 (0.032)	0.016 (0.045)	0.791 (0.032)		
Income	0.016 (0.001)	0.013 (0.001)	−0.021 (0.001)		
Commute	0.004 (0.003)	0.003 (0.005)	−0.028 (0.003)		
Gas_stations	−0.011 (0.002)	−0.009 (0.003)	−0.026 (0.002)		
Dealerships	−0.012 (0.004)	−0.015 (0.006)			
Refineries			0.316 (0.034)		
Interstates					26.608 (4.068)
Fleet installed base					1.415 (0.201)
Fleet installed base ∗ N_2					−1.226 (0.560)
ρ_1				0.024 (0.105)	
ρ_2				0.002 (0.111)	
Observations	34,713	34,713	34,713	34,713	34,713
Serial autocorrelation test (AR2 Z)	−0.761	−0.913			
AR2 p-value	0.447	0.361			
Overidentification test (Hansen J)	$\chi^2_7 = 7.725$	$\chi^2_{14} = 17.323$			
Overidentification p-value	0.357	0.239			
Zip FE	Y	Y	N	N	N
State FE	N	N	Y	Y	N
Year FE	Y	Y	Y	Y	N
State/Year FE	Y	Y	N	N	N
Zip time trends	Y	Y	N	N	N

6. Results

Model parameter estimates are reported in Table 4 and discussed in the sections that follow.

6.1. Flex-Fuel Demand

Flex-fuel demand model results are presented in the first two columns of Table 4. Before discussing the parameter estimates, I comment on diagnostic tests of the SGMM estimator. Recall from §5 that identification requires that no serial autocorrelation be present in the shocks (ε_1, ε_2). A test proposed by Arellano and Bond (1991) checks for serial autocorrelation in the GMM residuals by testing for second order autocorrelation in $\Delta\varepsilon$.[27] The test statistic is normally distributed under the null of no serial autocorrelation. The test Z scores are -0.761 and -0.913 for the consumer and fleet equations, respectively. Thus, the null hypotheses are not rejected at conventional (5%) significance levels, suggesting that use of the SGMM estimator is valid. Similarly, the Hansen J test statistics (which are distributed chi-squared with degrees of freedom equal to the number of overidentifying restrictions) are $\chi^2_7 = 7.725$ (consumer) and $\chi^2_{14} = 17.323$ (fleet). Again, the null of proper exclusion of the lagged values of N_1 and N_2 cannot be rejected at conventional significance levels.

Turning to the consumer FFV demand estimates, the key parameter (α_1) is the coefficient on the number of E85 retailers (N_1), which captures the demand side of the network effect. As anticipated, this coefficient is positive and significant. The parameter estimate implies that increasing the number of

[27] After transforming to first differences, testing for AR(1) in ε in the original estimation equation is equivalent to testing for AR(2) in $\Delta\varepsilon$. The test statistic is given by $\sum_{m,t} \varepsilon_{m,t}\varepsilon_{m,t-2}$.

Table 5 Comparison of α_1 and α_{21} Estimates Under Different Models

	OLS	OLS	OLS	SGMM
α_1	0.128	0.131	0.068	0.059
	(0.013)	(0.026)	(0.015)	(0.027)
$\alpha_2 1$	−0.643	0.088	0.091	0.064
	(0.023)	(0.027)	(0.021)	(0.031)
Observable controls	N	Y	Y	Y
Zip FE	N	Y	Y	Y
Year FE	N	N	Y	Y
State/year FE	N	N	Y	Y
Zip time trends	N	N	Y	Y

E85 retailers in a market by one increases the log-odds of FFV purchase (H_1) by 0.059. Interpretation is aided by translating this result into probability and unit demand measures. Noting that the probability of FFV purchase is related to the log-odds ratio by $\Pr[purchase] = \exp(H_1)/(1+\exp(H_1))$, I find the average marginal effect of adding one E85 retailer to a market is a 6.1% increase in the probability of FFV purchase.[28] As expected unit demand is given by $\hat{Q}_1 = P_1 \cdot \Pr[purchase]$, I find that adding one E85 retailer to a market on average increases annual unit sales by 0.87 FFVs. The structural demand model thus generates a more conservative prediction of the effect of E85 availability on FFV demand than is implied by the descriptive regressions of Table 3. As a further check that the SGMM estimator is producing sensible results, in Table 5 I compare the SGMM estimate of α_1 to a series of models that do not use instruments and use less stringent controls. As may be seen, the α_1 estimates are generally robust to the model specification. While the results are statistically consistent across specifications (i.e., the 95% confidence intervals overlap), the point estimates follow the expected pattern of correcting an upward bias through stronger controls and use of instruments.

The remaining coefficients (β_1) capture the effects of exogenous market-level observables (X_1) on consumer FFV demand. The demographic variables *rural*, *income*, and *commute* are included as controls because they may impact FFV demand via correlation with unobserved environmental preferences and fuel consumption patterns. Of these, only *income* is statistically significant. The positive sign of this coefficient is consistent with theories that environmental preferences correlate positively with income levels. *Gas_stations* controls for the availability of substitutes for E85, namely gasoline. The negative sign of the coefficient on *gas_stations* is sensible given that in markets where gasoline is readily available and less expensive, E85 and hence FFVs are relatively less attractive. *Dealers* controls for the availability of vehicles, the majority of which are substitutes for FFVs. The negative sign of this coefficient suggests that a greater number of substitutes is associated with lower FFV demand. For brevity, I suppress reporting and discussion of estimated fixed effects and time trend parameters.

Turning to the fleet flex-fuel demand estimates, the coefficient on N_1 (α_{21}) is also positive and significant, implying positive feedback in fleet flex-fuel demand and E85 retailer entry. The corresponding average marginal effect of adding one E85 retailer entry is a 6.5% increase in the probability of FFV purchase and an annual unit sales increase of 0.17 FFVs. Again the aggregate logit model generates a more conservative prediction of the effect of E85 availability on FFV demand than is implied by the descriptive regressions. The coefficient on private E85 stations is insignificant after controlling for market fixed effects, which is unsurprising since there is almost no within-market variation of this quantity. Table 5 also contains comparisons of the SGMM estimate of α_{21} to simpler models. In this case, inclusion of observable controls and market fixed effects has a major impact on model estimates, while additional corrections are negligible. I use the same exogenous market characteristics in the fleet demand equation as in the consumer equation (i.e., $X_2 = X_1$, where here *income* proxies for wages), and find similar dependence on these factors across the two equations.

6.2. Fuel Retailer Entry

The E85 market entry estimates are presented in columns three to five of Table 4. Entry model regressors are divided into categories corresponding to infrastructure cost shifters (X_3), variable profit shifters (X_4), and market size shifters (X_5). For infrastructure costs, the controls are state and year fixed effects, which I do not report for brevity. Here, state fixed effects capture variation in state-sponsored infrastructure incentives as well as regional differences in equipment and installation labor costs. Year effects capture changes in federal incentives and average E85 infrastructure costs. I include the cross-equation correlation parameters in the infrastructure cost column, as the econometric error term ε_3 also enters the profit function linearly. The estimated coefficients ρ_{13} and ρ_{23} are small and statistically insignificant. I interpret this result as indicating that the relevant codependence of the FFV demand and E85 entry systems is fully captured through observables and fixed effects in the model specification.

The variable profit specification also includes state and year effects to capture variation in excise tax

[28] Because N_1 is discrete, I report marginal effects as a proportional change rather than an elasticity. That is, I compute the average marginal effect as the sample average of $(\Pr(N_1 = 1) - \Pr(N_1 = 0))/\Pr(N_1 = 0)$. The marginal effect is constant (to one decimal place) over all observed transition values for N_1 (e.g., $0 \to 1, 1 \to 2$, etc.).

rates and average fuel price/cost margins. Observable shifters include demographics that may correlate with fuel type preferences (*rural*, *income*), fuel consumption rates (*commute*), wholesale ethanol prices (*refineries*), and substitute fuel prices (*gas_stations*). In general, the estimates are as expected. Profits are higher where wholesale ethanol prices are low (more ethanol plants) and substitute fuel prices are high (fewer gas stations). The fact that variable profits decrease in *commute* time potentially reflects that consumers with high fuel use costs are less likely to buy E85 because it is rarely priced lower on a \$/mile basis. That variable profits decrease in *income* is slightly counterintuitive, since higher income consumers presumably have higher willingness to pay for E85. One potential explanation is that the income effect is picking up higher input costs for E85, as high income neighborhoods tend to be far from ethanol refineries located in rural communities.

The market size coefficients are particularly interesting to interpret, as units have been normalized to consumer FFVs. The variable *interstates* is intended to capture E85 consumption by FFVs in long-range commuting patterns. The estimate implies that the contribution of an interstate highway to retailer profits is the same as that of 26.6 consumer FFVs in the installed base. Similarly, the coefficient on the fleet flex-fuel installed base (γ) parameter indicates that it takes 1.4 consumer FFVs to equal the profit contribution of one fleet FFV. This result is expected, given that fleets have many alternative fuel use incentives while consumers do not. In his study, Corts (2010) finds that it takes about twice as many *private* vehicles (which he defines as consumer plus corporate fleet FFVs) as *government* fleet FFVs to support one E85 station. A direct comparison of the results is difficult because of the different definitions, but I view this as a broadly consistent finding in that government fleets generally have stronger E85 consumption incentives than do corporate fleets. I also include an interaction term of the fleet installed base with the number of private E85 facilities in the market, anticipating that fleets with dedicated facilities will be "out of the market" and thus not contribute to consumption at retail E85 stations. The coefficient estimate of -1.2 is consistent with this scenario, suggesting that when a dedicated facility is present, most fleet E85 consumption is diverted there.

Interpretation of the entry model results is facilitated by computing the number of FFVs required to support a given number of E85 retailers, or the "entry thresholds" in the language of Bresnahan and Reiss (1991). The supply side of the network effect essentially operates through this mechanism: as the installed base of FFVs increases, the market becomes more profitable to serve, and at certain threshold values of the installed base, additional entry becomes feasible. I report the entry thresholds in Table 6. The threshold values are computed from the entry model estimates using the formula: $S^*_{N_1} = (\exp(\hat{\psi}' X_4)/(\hat{\phi}' X_5))(N_1+1)^2$. I compute entry thresholds at sample average values of the controls, $\bar{X}_4$ and $\bar{X}_5$. Given significant right-skew in the distribution of the predicted thresholds, I consider the median to be the better measure of central tendency. By this measure, the model predicts that under typical conditions at least 299 consumer FFVs are required to support a single E85 retailer. My estimates appear broadly in line with those obtained independently by other research methods. A joint publication by the EPA and the Departments of Transportation and Energy (U.S. DOT, DOE, and the EPA 2002) reports that roughly 200 FFVs are required to support an E85 retailer. Corts (2010) estimates a series of models (linear, Poisson, Tobit) that suggest between 320 and 560 private FFVs are needed to support an E85 station.

Table 6 E85 Market Entry Thresholds in Consumer FFV Units

E85 retailers	Mean	Std dev	Median
1	402	302	299
2	904	679	673
3	1,607	1,207	1,196
4	2,510	1,886	1,868
5	3,615	2,716	2,690

The E85 entry model entails several strong assumptions, including that market outcomes are independent and that firm competition is accurately represented as a Cournot output game. In Web Appendix E, I develop robustness checks that relax these assumptions and find that my results remain highly consistent with those reported in Tables 4 and 6.

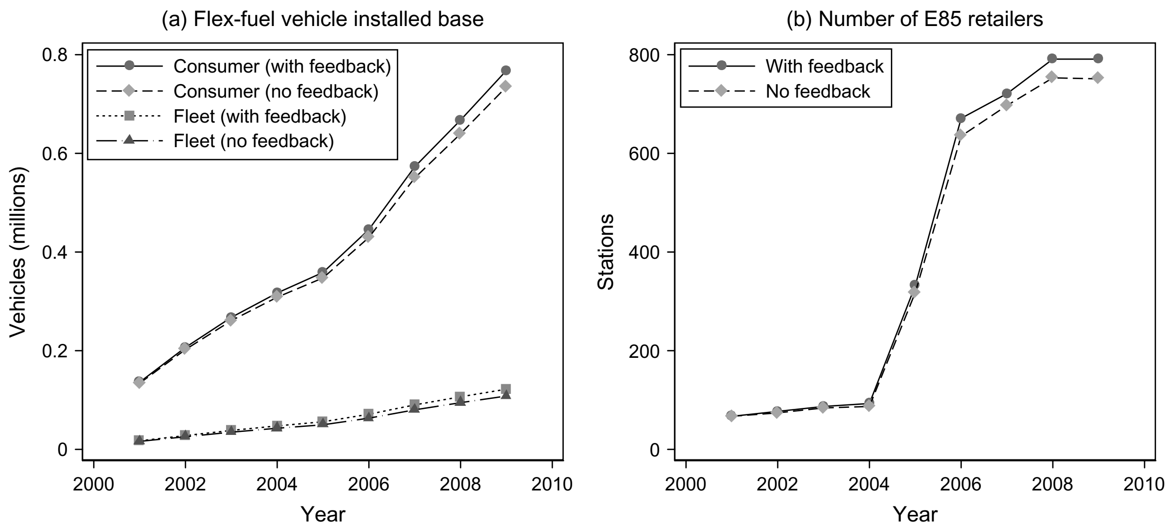
7. Applications and Strategic Implications

In this section, I demonstrate application of the results. The first application quantifies the long-run effect of positive feedback across a sample of heterogeneous markets, which cannot be directly inferred from the model estimates. The second experiment explores promotional strategies that firms may implement to accelerate adoption of the ethanol fuel standard, and thereby generate additional demand for their products. For this experiment, I take the perspective of an FFV manufacturer and evaluate the effect of offering promotional subsidies to fuel retailers to enter the E85 market.

7.1. Quantifying the Network Effect

To quantify the long-run impact of the network effect, I first simulate market outcomes assuming that the model estimates are the true parameters governing

Figure 1 Quantifying the Long-Run Impact of the Network Effect



the data generating process (the "with feedback" condition). Then, the simulation is repeated under the assumption that neither consumers nor fleets have utility for E85 (i.e., for the "no feedback" condition, I set $\alpha_1 = \alpha_{21} = 0$). Simulation of market outcomes requires the following series of steps. First, for each market observation, I draw market shocks from the distribution $N(0, \hat{\Sigma})$. Next, I solve the model for the equilibrium values of Q_1^*, Q_2^*, and N_1^* (as during the entry model estimation) period by period, updating the installed base sequentially. Third, I repeat the first two steps multiple (30) times and average the result to obtain the expected value of the market outcomes.

I summarize the counterfactual graphically in Figures 1(a) and 1(b). In these plots, the installed bases and number of E85 retailers are aggregated across the sample markets. At any point in time, the cumulative influence of the network effect is given by the difference between the "with feedback" and "no feedback" curves. In the final period (2009), the network effect accounts for 5.3% of the predicted number of E85 retailers. Similarly, the network effect accounts for 5.6% of the total FFV installed base. Indirect network effects of this size are economically material, but modest compared to those observed in high-tech product markets. For example, Nair et al. (2004) find that indirect network effects in PDA hardware/software adoption account for 22% of the installed base of PDAs over a period roughly one-third of the length of my study.

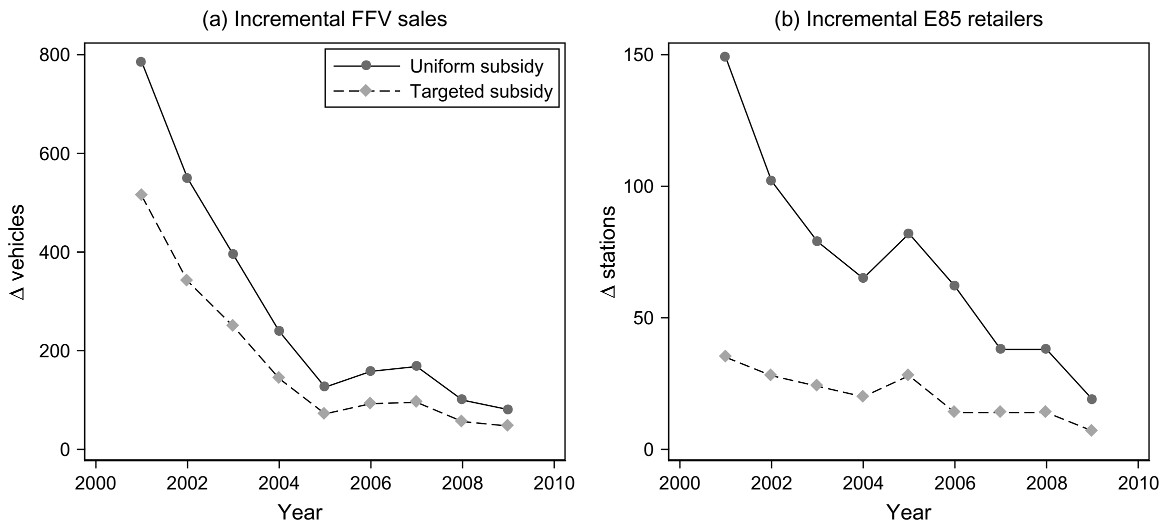
7.2. FFV Manufacturer Subsidy of E85 Entry

I assess two joint marketing arrangements in which FFV manufacturers provide subsidies to fuel retailers to encourage E85 market entry and thereby increase FFV sales. The first policy considers a fixed-rate subsidy that reduces the average fixed cost of E85 market entry by 25% for all potential entrants in the first period (2001). The effectiveness of the policy is measured by comparing outcomes to the baseline "with feedback" simulation used to quantify the long-run network effect: The benefit of the policy is the incremental number of FFV sales generated by the intervention, and the cost is measured by the number of stations that use the subsidy (i.e., the number of new entrants in 2001). The second policy involves the same timing and level of subsidy, but allows the vehicle manufacturer to target specific markets with higher anticipated rates of return. Attractive markets are characterized by two factors, i.e., higher levels of expected FFV demand, and higher entry thresholds. The latter follows because it is not cost effective to subsidize a market where entry is likely to occur naturally anyway. I evaluate a simple targeting rule: markets that are offered the subsidy must be in the top quartile of expected FFV demand ($E[Q_1]$) and the top quartile of entry thresholds. By this criteria, 17% (656/3,857) of markets are eligible for the targeted subsidy.

Figures 2(a) and 2(b) summarize the effect of these counterfactual policies. Plotted values reflect the incremental effect of the policy on FFV unit sales and the number of E85 retailers relative to the baseline "with feedback" simulation in §7.1. As expected, the uniform policy generates more incremental FFV unit sales, but at a heavy cost. Whereas the targeted policy leads to 43 E85 retailers entering in 2001, the uniform policy induces 157 new entrants. As time passes, the effects of both policies diminish relative to the baseline simulation because many subsidized markets eventually experience entry naturally. However, the rate at which the intervention effect declines is more extreme for the uniform policy.

Some back-of-the-envelope calculations can shed additional light on the cost/benefit proposition of the

Figure 2 Subsidy Counterfactual



two policies. In Table 7, I first calculate the costs of the two policies in "E85 infrastructure" units by multiplying the number of entrants induced by the policy by 25%. Next I calculate the present value of incremental FFV sales (in 2001 vehicle units) for the two policies, assuming a 5% discount rate (and no terminal value). Benefit to cost ratios derived from these figures are sufficient to compare the relative attractiveness of the two policies. The advantage of targeting is readily apparent: the value proposition of the targeted policy is more than twice that of the uniform policy. Of course, computing the net benefit in dollar terms requires knowledge of the manufacturer margin per FFV and average E85 infrastructure costs. However, only the ratio of these quantities is needed to define a break-even "hurdle rate" for a policy: a policy will be profitable provided its benefit/cost ratio exceeds the ratio of E85 infrastructure costs to FFV margins.[29] As an example, the targeted policy will be profitable provided the ratio of E85 infrastructure costs to FFV margins is less than 136. What constitutes a realistic hurdle rate? Secondary sources provide some guidance. The discussion in §2.2 suggests that the average cost of E85 infrastructure investments lies in the range \$10,000–\$50,000; financial reporting on automakers suggests that average vehicle margins are in the low thousands.[30] Thus, a fairly conservative estimate of this ratio would be around \$50,000/\$500 = 100 while a liberal one might be \$10,000/\$2,000 = 5. In the conservative case, the targeted policy would be profitable whereas the uniform policy would not. This simple experiment demonstrates that the network effect can be harnessed to improve vehicle manufacturer profitability, and that a targeted incentive policy incorporating knowledge of local market conditions offers a considerable boost to returns on investment.

Table 7 Subsidy Cost/Benefit Analysis

Policy	Uniform	Targeted
Subsidized stations	157	43
Cost (E85 infrastructure units)	39.3	10.8
Benefit (2001 FFVs)	2,351.5	1,464.0
Benefit/cost ratio	60.0	136.19

8. Conclusion

This paper presents evidence of indirect network effects in the demand for ethanol-compatible vehicles and the supply of ethanol fuel. In contrast to most studies of indirect network effects, the feedback mechanism investigated here is spatial in nature and operates at a highly localized level. To identify these local effects, I estimate a model of consumer FFV demand, fleet FFV demand, and E85 retailer market entry using a zip code panel data set that incorporates the entire population of FFV registrations and E85 market entry events in four states over nine years. The framework extends the Bresnahan and Reiss (1990, 1991, 1994) models of competitive entry to incorporate an endogenous market size. I develop a new simulation-based estimator that accommodates simultaneity and correlated shocks across the model equations, without requiring the use of instruments for the entry model. The model estimates indicate a network effect with

[29] To see this, let (w_m, q_m) be FFV (margin, incremental sales) and (w_r, q_r) be E85 (average infrastructure costs, incremental retailers) and s be the subsidy rate. A policy is profitable if $w_m q_m \geq s w_r q_r$, and thus if $q_m / s q_r \geq w_r / w_m$.

[30] Margins on new vehicles sold in the United States vary greatly by manufacturer and model. Margins on pickup trucks are typically higher, often by more than \$5,000, compared to hundreds of dollars on most cars. See http://money.msn.com/top-stocks/post.aspx?post=620b56b9-a45d-4036-bd1b-9b66461d978d.

both statistical and economic significance, but moderate size compared to those found in high tech product markets. As an application, I demonstrate that the network effect may be leveraged to improve FFV manufacturer profitability through targeted subsidies of E85 market entry.

There are some limitations to the analysis and therefore several possibilities for extension. Perhaps the most obvious limitation is the restrictive treatment of dynamics. In particular, forward-looking behavior by consumers, fleets, and fuel retailers is not formally modeled. However, abstraction from dynamic considerations in the FFV demand system is unlikely to be a major concern for the empirical estimates due to: (a) the absence of incremental sunk costs to acquire FFV technology and no irreversible commitment to E85 use, (b) the high persistence of E85 availability levels (implying myopic conditioning on current levels closely approximates rational future expectations), and (c) the inclusion of rich fixed effects in the model that can, to some extent, capture expectations flexibly in a nonparametric fashion. The treatment of dynamics in the E85 supply model is likely a more serious limitation. While the model admits a form of dynamically-evolving firm heterogeneity by distinguishing between incumbents and potential entrants, it captures firms' expected discounted profits as a reduced form rather than through a recursive computation explicitly incorporating firm expectations. As previously noted, absent all sunk costs, these procedures would yield equivalent decision rules, but the fact that some sunk costs do exist means there is some potential for specification-related bias. However, inasmuch as arguments (b) and (c) above also apply to the E85 entry model, and to the extent that sunk costs are modest in this context relative to most empirical studies of market entry, I view this approach as reasonable for the purpose of measuring the network effects of interest. Nevertheless, this issue is worthy of further attention. The paper could be extended by relaxing the assumption of identical firms in the entry model. In this vein, exploring multimarket contact and scale economies in ethanol distribution are also potentially interesting topics of study. I leave these extensions to future work.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mksc.2014.0881.

Acknowledgments

The author thanks the thesis committee (Harikesh Nair, Wes Hartmann, Peter Reiss, Han Hong, and Jon Levin) for their guidance and comments. The author also thanks the editor, associate editor, and two anonymous reviewers, as well as Seenu Srinivasan, Jim Lattin, Sridhar Narayanan, Michaela Draganska and Seethu Seetharaman for their extensive comments. The author gratefully acknowledges comments from seminar participants at the 9th Invitational Choice Symposium, Marketing Science, Cornell University, the University of Michigan, the University of Chicago, Northwestern University, the University of California, Berkeley, the University of Maryland, Columbia University, New York University, and Washington University in St. Louis. The author thanks Margot Melendez at the National Renewable Energy Laboratory (NREL) for assistance in obtaining E85 station data, and Tim Gerlach at the American Lung Association of the Upper Midwest for many helpful discussions on the E85 supply chain. Generous financial support from the Stanford Graduate School of Business (GSB) and the NET Institute [grant 10-20] is also gratefully acknowledged. All errors are the author's.

References

Anderson S (2011) The demand for ethanol as a gasoline substitute. *J. Environ. Econom. Management* 63(2):151–168.

Anderson S, Sallee J (2011) Using loopholes to reveal the marginal cost of regulation: The case of fuel-economy standards. *Amer. Econom. Rev.* 101(4):1375–1409.

Arellano M, Bond S (1991) Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Rev. Econom. Stud.* 58(2):277–297.

Basu A, Mazumdar T, Raj S (2003) Indirect network externality effects on product attributes. *Marketing Sci.* 22(2):209–221.

Berry S (1994) Estimating discrete-choice models of product differentiation. *RAND J. Econom.* 25(2):242–262.

Berry S, Waldfogel J (1999) Free entry and social inefficiency in radio broadcasting. *RAND J. Econom.* 30(3):397–420.

Blundell R, Bond S (1998) Initial conditions and moment restrictions in dynamic panel data models. *J. Econometrics* 87(1):115–143.

Bresnahan T, Reiss P (1990) Entry in monopoly markets. *Rev. Econom. Stud.* 57(4):531–553.

Bresnahan T, Reiss P (1991) Entry and competition in concentrated markets. *J. Political Econom.* 99(51):977–1009.

Bresnahan T, Reiss P (1994) Measuring the importance of sunk costs. *Annales d'Economie et de Statistique* 34:181–217.

Chou C, Shy O (1990) Network effects without network externalities. *Internat. J. Indust. Organ.* 8(2):259–270.

Church J, Gandal N (1992) Network effects, software provision, and standardization. *J. Indust. Econom.* 40(1):85–103.

Church J, Gandal N (1993) Complementary network externalities and network adoption. *Internat. J. Indust. Organ.* 11(2):239–260.

Clements M, Ohashi H (2005) Indirect network effects and the product cycle: Video games in the U.S., 1994–2002. *J. Indust. Econom.* 53(4):515–542.

Corts K (2010) Building out alternative fuel retail infrastructure: Government fleet spillovers in E85. *J. Econom. Management Strategy* 59(3):219–234.

Davis S, Diegel S (2006) Transportation energy data book: Edition 25. Technical report, ORNL/TM-2007/070, Oak Ridge National Laboratory (ORNL), Oak Ridge, TN.

Dixit A, Pindyck R (1994) *Investment Under Uncertainty* (Princeton University Press, Princeton, NJ).

Draganska M, Mazzeo M, Seim K (2009) Beyond plain vanilla: Modeling joint product assortment and pricing decisions. *Quant. Marketing Econom.* 7(2):105–146.

Dranove D, Gandal N (2003) The DVD-vs.-DIVX standard war: Empirical evidence of network effects and preannouncement effects. *J. Econom. Management Strategy* 12(3):363–386.

Dubé J, Hitsch G, Chintagunta P (2010) Tipping and concentration in markets with indirect network effects. *Marketing Sci.* 29(2):216–249.

Eppli M, Benjamin J (1994) The evolution of shopping center research: A review and analysis. *J. Real Estate Res.* 9(1):5–32.

Farrell J, Saloner G (1985) Standardization, compatibility, and innovation. *RAND J. Econom.* 16(1):70–83.

Gandal N, Kende M, Rob R (2000) The dynamics of technological adoption in hardware/software systems: The case of compact disc players. *RAND J. Econom.* 31(1):43–61.

Hansen L (1982) Large sample properties of generalized method of moments estimators. *Econometrica* 50(4):1029–1054.

Harris Interactive (2006) U.S. consumer awareness of E85 flex-fuel vehicles high, buying interest split about 50:50. Accessed April 2013, www.greencarcongress.com.

Karaca-Mandic P (2011) Role of complementarities in technology adoption: The case of DVD players. *Quant. Marketing Econom.* 9(2):179–210.

Katz M, Shapiro C (1994) Systems competition and network effects. *J. Econom. Perspectives* 8(2):93–115.

Laffont J, Ossard H, Vuong Q (1995) Econometrics of first-price auctions. *Econometrica* 63(4):953–980.

Lee RS (2013) Vertical integration and exclusivity in platform and two-sided markets. *Amer. Econom. Rev.* 103(7):2960–3000.

Liu H (2010) Dynamics of pricing in the video game console market: Skimming or penetration? *J. Marketing Res.* 47(3):428–443.

Maritz Automotive Research Group (2006) Four out of five consumers would consider flex-fuel vehicle for next purchase. Accessed April 2013, www.maritz.com.

Nair H, Chintagunta P, Dubé J (2004) Empirical analysis of indirect network effects in the market for personal digital assistants. *Quant. Marketing Econom.* 2(1):23–58.

National Renewable Energy Laboratory (2008) Cost of adding E85 fueling capability to existing gasoline stations: NREL survey and literature search. Accessed November 2012, www.afdc.energy.gov.

Ohashi H (2003) The role of network effects in the U.S. VCR market, 1978–1986. *J. Econom. Management Strategy* 12(4):447–494.

Park S (2004) Quantitative analysis of network externalities in competing technologies: The VCR case. *Rev. Econom. Statist.* 86(4):937–945.

Petrolia DR, Bhattacharjee S, Hudson D, Herndon CW (2010) Do Americans want ethanol? A comparative contingent-valuation study of willingness to pay for E10 and E85. *Energy Econom.* 32(1):121–128.

Pettigrew H, Gart J, Thomas D (1986) The bias and higher cumulants of the logarithm of a binomial variate. *Biometrika* 73(2): 425–435.

Phoenix Automotive (2006) FFV & E85 perceptions. Accessed December 2012, www.cleanairchoice.org.

Phoenix Automotive (2007) E85 everywhere: Minnesota consumer perceptions. Accessed December 2012, www.cleanairchoice.org.

Reid K (2006) Alcohol at the pump. *Natl. Petroleum News* 97(9): 40–44.

Rysman M (2004) Competition between networks: A study of the market for yellow pages. *Rev. Econom. Stud.* 71(2):483–512.

Shankar V, Bayus B (2003) Network effects and competition: An empirical analysis of the home video game industry. *Strategic Management J.* 24(4):375–384.

U.S. DOT, DOE, and the EPA (2002) Effects of the alternative motor fuels act CAFE incentives policy. Technical report, United States Congress, Washington, DC.

U.S. EIA (2011) Ethanol blend wall: Are we there yet? Accessed June 2013, www.eia.gov.

U.S. EIA (2012) On-road AFVs made available by year. Accessed July 2013, www.afdc.energy.gov.

Vitorino M (2012) Empirical entry games with complementarities: An application to the shopping center industry. *J. Marketing Res.* 49(2):175–191.