The Effects of Product Line Breadth: Evidence from the Automotive Industry. Online Appendix.

Antonio Moreno
Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston, IL 60208
a-morenogarcia@kellogg.northwestern.edu

Christian Terwiesch
Operations and Information Management, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104
terwiesch@wharton.upenn.edu

History: First version, April 2013. Second version, July 2014. Third version, August 2015. Fourth version, December 2015.

A. Robustness Analysis: Single-Equation Estimation Models

The estimation of the system of simultaneous equations using 3SLS is consistent and asymptotically more efficient than single-equation estimation when the model is correctly specified. However, single-equation models can provide more reliable results if the system of equations is not correctly specified. Rossi (2014) has recently noted some potential problems associated with the use of instrumental variables in marketing. In order to make sure the support (or lack thereof) of our hypothesis is not a consequence of misspecification of the model or our specific choice of instruments, we conduct a series of single-equation analyses to evaluate the robustness of our findings, including estimation by OLS and 2SLS.

We use the analysis of the effect of product line breadth on incentives as an example of our single-equation estimation approach. The single equation for the make-level incentives is the following:

$$\begin{aligned} DISCMAKE_{jt} =& \delta''_j + \beta''_1 PLB_{jt} + \beta''_2 MPGSPAN_{jt} + \beta''_3 GASVOL_t + \beta''_4 MPGSPAN_{jt} \times GASVOL_t \\ & + C``_{jt} + \sum_{k=1}^{K} \eta''_k INVMAKE_{jt-k} + \omega''_t + \epsilon''_{2jt} \end{aligned} \tag{1}$$

where j denotes make and t denotes month. In this single equation, make fixed effects δ''_j will control for any omitted make-level time-invariant factors. For example, the fact that Chevrolet is an American company is accounted for by the corresponding automaker fixed effects. Time controls ω_t'' account for temporal trends in the market. As in the system estimation approach, the specifications of this family differ in the number and nature of the variables that we include as controls, which are denoted by $C``_{jt}$. In any case, ϵ_{2jt}'' absorbs the part of the make discount that is not explained by the variables included in the model.

A potential concern with estimating Equation 1 (and some of the other single equation models we describe below) using OLS is endogeneity bias. Given that it is the firm's decision to expand or reduce its product line, we might be worried this decision could be correlated with some unobserved variables that explain some of the variation in incentives the firm offers, or that this variable could be simultaneously determined with other important variables of the system.

Our fixed effects somewhat mitigate this concern because they take care of any unobserved variables that are time-invariant, but we may still have time-variant unobserved variables that might be correlated with product line breadth. For example, there might be a temporal demand shock for one brand that might increase market share and, at the same time, induce the firm to increase its product line. Note that this is unlikely, since extending or reducing the product line is a decision typically made way before any temporal shocks in demand can affect market shares. Automotive manufacturers often follow multiyear plans when they decide on their product introduction and discontinuation schedule. If firms fully adhered to these predetermined schedules, the variation in their product line breadth could almost be seen as exogenous when it comes to studying the discounts they provide. (In the extreme, one could almost think of it as a form of quasi-experiment, where the pre-established schedule determines the product line breadth a long time in advance.)

Nevertheless, deviations from the plan can be endogenous and it is important to take them into account. Therefore, we propose an alternative instrumental variable estimation of the effect of product line breadth on make-level incentives using 2SLS to account for the endogeneity of product line breadth. We use the number of products manufactured by the rest of the makes as an instrument. Similar instruments (together with many other model-level instruments that we are not able to use at the make level) have been used in previous research in the automotive industry (see Berry et al. 1995). The underlying assumption of this analysis is that there will be some correlation between the number of products of the rest of the makes and the product line breadth of a make j, but that the number of products of the rest of the makes will not directly affect the incentives make j provides.

We proceed in an analogous way to estimate single-equation specifications of the market share and inventory equations. In all cases, we estimate the coefficients using OLS and 2SLS, and a variety of specifications. The results of the estimation of the single-equation specifications at the

Table 5 Effect of Product Line Breadth on Market Shares: Make-level Single-equation Estimation

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
PLB	0.00106***	0.00104***	0.00109***	0.00103***	0.00139***	0.00137***	0.00142***	0.00137***
	(0.000187)	(0.000192)	(0.000188)	(0.000187)	(0.000180)	(0.000184)	(0.000181)	(0.000180)
$MPGSPAN$		-0.00146**	-0.00144**	-0.00137**		-0.00146**	-0.00144**	-0.00137**
		(0.000618)	(0.000608)	(0.000610)		(0.000608)	(0.000597)	(0.000599)
$MPGSPAN \times GASVOL$		7.73e-05**	8.73e-05**	9.48e-05***		7.18e-05**	8.19e-05**	8.92e-05**
		(3.56e-05)	(3.60e-05)	(3.61e-05)		(3.51e-05)	(3.55e-05)	(3.55e-05)
$MEDMPGMAKE$			-0.00196***	-0.00200***			-0.00198***	-0.00202***
			(0.000215)	(0.000216)			(0.000212)	(0.000213)
$MEDMPGMAKE \times GAS$			5.77e-06***	6.16e-06***			5.79e-06***	6.17e-06***
			(6.02e-07)	(6.22e-07)			(5.94e-07)	(6.12e-07)
$DISCMAKE$				7.66e-07***				7.37e-07***
				(1.18e-07)				(1.16e-07)
$FIXED\ EFFECTS$	Make	Make	Make	Make	Make	Make	Make	Make
$TIME\ CONTROLS$	Month-Year	Month-Year	Month-Year	Month-Year	Month-Year	Month-Year	Month-Year	Month-Year
$OTHER\ CONTROLS$	No	Yes^+	Yes^+	Yes^+	No	Yes^+	Yes^+	Yes^+
Observations	2,994	2,962	2,962	2,956	2,994	2,962	2,962	2,956
R-squared	0.966	0.966	0.967	0.967	0.015	0.018	0.043	0.049

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$
$^+$ indicates $GASPRICE$, $GASVOL$ as additional controls
Columns 5-8 use the total number of models of other makes as an IV.
The Durbin-Wu-Hausman test rejects the hypothesis of exogeneity (e.g., in Column 8, p<0.0001) and the 2SLS estimates are preferred.
The Angrist-Pischke multivariate F test of excluded instruments supports does not suggest weak instruments (e.g., in column 8, p<0.0001)

make level are shown in Table 5 (market shares), Table 6 (discounts) and Table 7 (inventories). In all cases, the results are qualitatively consistent with results obtained using 3SLS estimation.

As we have done above for the make-level single-equation estimation, we describe the model-level single-equation estimation approach for the effect of product line breadth on incentives. The discussion for the other equations of interest of the system of simultaneous equations described above would be analogous. Let i denote a model marketed by make j. The single equation for the model level incentives is the following:

$$\begin{aligned} DISCMODEL_{it} =& \delta_i''' + \beta_1''' PLB_{jt} + \beta_2''' MPGSPAN_{jt} + \beta_3''' GASVOL_t + \beta_4''' MPGSPAN_{jt} \times GASVOL_t \\ &+ \beta_5''' MKTSHR_{it} + \sum_{k=1}^{K} \eta_k''' INVMODEL_{it-k} + C'''_{it} + \omega_{st}''' + \varepsilon_{3it}''' \end{aligned} \tag{2}$$

In this single equation, the dependent variable $DISCMODEL_{it}$ is the average discount given for model i in month t (i.e., the sum of all the money spent on discounts in month t for model i over the number of model i vehicles sold in month t). On the right-hand side, model fixed effects δ_i''' will control for any omitted model-level time-invariant factors. In a model-level specification, a richer set of controls can be used, compared with the make-level specification. For example, we include time-segment controls ω_{st}''' that account for different discount patterns for vehicles belonging to different segments. As in the system estimation approach, the specifications of this family differ in the number and nature of the variables we include as controls, denoted by C'''. In any case, ϵ_{3it}''' absorbs the part of the model discount that is not explained by the variables included in the model.

Obviously, not all observed variation in discount levels indicates mismatch costs. On one hand, it is well known that some companies have a higher inherent tendency to use discounts. Our model fixed effects account for that. On the other hand, discounts can be the firm's reaction to a different competitive scenario. For example, if the competitors of a model give discounts, a firm might be forced to use discounts even in a situation where the production decision was "ex-ante right." A model that fully characterizes the equilibrium behavior of discounts is beyond the objectives of this paper. However, in order to ensure that our results do not follow from these competitive aspects, we control for the average level of discounts offered by the competitors of a model (i.e., same segment and luxury level) in our set of controls C'''_{it}. Together with our segment–time interactions, this captures the evolution of the market for the different segments.

Another potential concern is that there might be endogeneity. Product line decisions might be correlated with unobserved variables that explain part of the discount behavior, or they could be simultaneously determined with other important variables of the system. We use an exhaustive control strategy to attenuate this concern. Furthermore, it is important to note that product line decisions are made long before any unaccounted temporal shock that could affect discounts is realized. As noted above in the discussion of the make-level analysis, automotive manufacturers often follow multiyear plans when they decide on their product introduction and discontinuation schedule, which can result in product line breadth behaving as exogenous when it comes to studying model discounts.

In order to make sure our results are not driven by the aforementioned forms of endogeneity (omitted variable and simultaneity bias), we propose an alternative instrumental variable estimation of the effect of product line breadth on model-level incentives using 2SLS to account for the endogeneity of product line breadth. We use a set of instruments based on product characteristics of other makes' models. This is an approach Berry et al. (1995) use, which considers product characteristics as exogenous. To be specific, we include the sum of the $SIZE$, $HPWT$, and MPG of the rest of the models of the same make; the sum of the $SIZE$, $HPWT$, and MPG of the models of the rest of the makes; and the number of other makes' models.

We proceed in an analogous way to estimate single-equation specifications of the market share and inventory equations. In all cases, we estimate the coefficients using OLS and 2SLS and a variety of specifications. The results of the estimation of the single-equation specifications at the model level are shown in Table 9 (estimated production costs), Table 10 (discounts) and Table 11 (inventories). In all cases, the results are qualitatively consistent with the results obtained using 3SLS estimation.

Table 6 Effect of Product Line Breadth on Mismatch Costs: Discounts. Make-level Single-equation Estimation

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
PLB	57.72***	66.25***	67.50***	65.55***	55.56***	65.83***	67.43***	65.02***
	(17.54)	(17.36)	(17.38)	(17.78)	(17.18)	(16.44)	(17.00)	(17.39)
$MPGSPAN$		-132.8**	-103.9*	-98.63		-132.8**	-103.9*	-98.61
		(63.83)	(61.54)	(66.80)		(58.77)	(60.21)	(65.31)
$MPGSPAN \times GASVOL$		-9.296***	-8.997***	-8.980***		-9.289***	-8.996***	-8.973***
		(3.333)	(3.398)	(3.404)		(3.005)	(3.324)	(3.327)
$INVMAKE_{t-1}$				-1.197				-1.196
				(0.955)				(0.934)
$INVMAKE_{t-2}$				1.871*				1.872*
				(1.078)				(1.054)
$INVMAKE_{t-3}$				2.174**				2.174**
				(1.105)				(1.080)
$MEDMPGMAKE$			49.13*	38.35			49.13*	38.39
			(25.92)	(27.05)			(25.35)	(26.43)
$MEDMPGMAKE \times GAS$			-0.504***	-0.454***			-0.504***	-0.454***
			(0.0870)	(0.0902)			(0.0851)	(0.0882)
$FIXED\ EFFECTS$	Make	Make	Make	Make	Make	Make	Make	Make
$TIME\ CONTROLS$	Month-Year	Month-Year	Month-Year	Month-Year	Month-Year	Month-Year	Month-Year	Month-Year
$ADDIT.\ CONTROLS$	No	Yes^{o}	Yes^{o}	Yes^{o}	No	Yes^{o}	Yes^{o}	Yes^{o}
Observations	2,988	2,956	2,956	2,857	2,988	2,956	2,956	2,857
R-squared	0.735	0.755	0.760	0.758	0.216	0.218	0.232	0.202

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$
^{o} indicates $GASPRICE$, $GASVOL$
Columns 5-8 use the total number of models of other makes as an IV.
The Durbin-Wu-Hausman test of endogeneity does not allow to reject the hypothesis that PLB is exogenous (e.g., in column 8, p=0.94)
The Angrist-Pischke multivariate F test of excluded instruments supports does not suggest weak instruments (e.g., in column 8, p <0.0001)

Table 7 Effect of Product Line Breadth on Mismatch Costs: Inventories. Make-level Single-equation Estimation

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
PLB	3.116***	3.097***	3.100***	3.183***	3.144***	3.159***	3.166***	3.262***
	(0.448)	(0.450)	(0.450)	(0.448)	(0.441)	(0.441)	(0.441)	(0.438)
$MPGSPAN$		-3.206**	-3.006*	-3.062**		-3.206**	-3.005**	-3.062**
		(1.521)	(1.534)	(1.526)		(1.488)	(1.500)	(1.492)
$MPGSPAN \times GASVOL$		0.0736	0.0753	0.0562		0.0726	0.0742	0.0549
		(0.0868)	(0.0864)	(0.0855)		(0.0849)	(0.0845)	(0.0837)
$DISCMAKE$				-0.00135***				-0.00136***
				(0.000490)				(0.000479)
$FIXED\ EFFECTS$	Make	Make	Make	Make	Make	Make	Make	Make
$TIME\ CONTROLS$	Month-Year	Month-Year	Month-Year	Month-Year	Month-Year	Month-Year	Month-Year	Month-Year
$ADDIT.\ CONTROLS$	No	Yes^{o}	Yes^{a,o}	Yes^{a,o}	No	Yes^{o}	Yes^{a,o}	Yes^{a,o}
Observations	2,962	2,930	2,930	2,925	2,962	2,930	2,930	2,925
R-squared	0.687	0.697	0.697	0.699	0.392	0.401	0.402	0.408

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$
^{a} indicates $MEDMPGMAKE$, $MEDMPGMAKExGASPRICE$
^{o} indicates $GASPRICE$, $GASVOL$
Columns 5-8 use the total number of models of other makes as an IV.
The Durbin-Wu-Hausman test of endogeneity does not allow to reject the hypothesis that PLB is exogenous (e.g., in column 8, p= 0.11)
The Angrist-Pischke multivariate F test of excluded instruments does not suggest weak instruments (e.g., in column 8, p <0.0001)

Table 8 Effect of Product Line Breadth on List Prices. Model-level Single-equation Estimation

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
PLB	59.06***	49.43***	42.41***	49.54***	83.08***	73.00***	64.57***	71.46***
	(14.67)	(14.36)	(14.10)	(14.39)	(14.57)	(14.19)	(13.95)	(14.22)
$PLATVOLUME$				-0.000397***				-0.000408***
				(0.000117)				(0.000115)
$FIXED\ EFFECTS$	Model	Model	Model	Model	Model	Model	Model	Model
$TIME\ CONTROLS$	Seg.xYear	Seg.xYear	Seg.xYear	Seg.xYear	Seg.xYear	Seg.xYear	Seg.xYear	Seg.xYear
$ADDIT.\ CONTROLS$	No	+	+, ^{o}	+, ^{o}	No	+	+, ^{o}	+, ^{o}
Observations	18,166	17,683	17,683	17,683	17,683	17,683	17,683	17,683
R-squared	0.985	0.986	0.986	0.986	0.663	0.647	0.657	0.661

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$
$^{+}$ indicates $SIZE$, $HPWT$, MPG
^{o} indicates $LAUNCHED$, $PHASEDOUT$, NEW_DESIGN, AGE
Columns 5-8 use the sum of characteristics ($SIZE$, $HPWT$ and MPG) of other models of the make,
the sum of the characteristics of the models of other makes and the total number of products of other makes as IVs.
The Durbin-Wu-Hausman test of endogeneity rejects the hypothesis that PLB is exogenous (e.g., in column 8, p<0.0001)
The Angrist-Pischke multivariate F test of excluded instruments does not suggest weak instruments (e.g., in column 8, p <0.0001)

Table 9 Effect of Product Line Breadth on Estimated Production Costs. Model-level Single-equation Estimation

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
PLB	150.4***	146.8***	150.8***	167.1***	148.1***	146.1***	150.9***	166.3***
	(34.31)	(34.73)	(35.40)	(36.10)	(33.06)	(33.33)	(33.87)	(34.46)
PLATVOLUME				-0.000913***				-0.000912***
				(0.000119)				(0.000117)
FIXED EFFECTS	Model	Model	Model	Model	Model	Model	Model	Model
TIME CONTROLS	Seg.xYear	Seg.xYear	Seg.xYear	Seg.xYear	Seg.xYear	Seg.xYear	Seg.xYear	Seg.xYear
ADDIT. CONTROLS	No	$+$	$+$, ^{o}	$+$, ^{o}	No	$+$	$+$, ^{o}	$+$, ^{o}
Observations	18,166	17,683	17,683	17,683	17,683	17,683	17,683	17,683
R-squared	0.915	0.914	0.914	0.914	0.572	0.571	0.573	0.573

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$
$^+$ indicates $SIZE$, $HPWT$, MPG
^{o} indicates $LAUNCHED$, $PHASEDOUT$, NEW_DESIGN, AGE
Columns 5-8 use the sum of characteristics ($SIZE$, $HPWT$ and MPG) of other models of the make,
the sum of the characteristics of the models of other makes and the total number of products of other makes as IVs.
The Durbin-Wu-Hausman test of endogeneity rejects the hypothesis that PLB is exogenous (e.g., in column 8, p<0.0001)
The Angrist-Pischke multivariate F test of excluded instruments does not suggest weak instruments (e.g., in column 8, p <0.0001)

Table 10 Effect of Product Line Breadth on Mismatch Costs: Discounts. Model-level Single-equation Estimation

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
PLB	106.1***	129.2***	126.8***	144.1***	105.6***	128.6***	129.2***	145.6***
	(10.30)	(10.74)	(10.67)	(10.84)	(10.16)	(10.54)	(10.49)	(10.64)
MPGSPAN			-281.0***	-318.9***			-287.6***	-322.0***
			(38.57)	(40.04)			(38.40)	(39.83)
$MPGSPAN \times GASVOL$			-6.872***	-4.734**			-6.723***	-4.655**
			(2.144)	(2.145)			(2.160)	(2.152)
COMPDISC		0.389***	0.326***	0.338***		0.372***	0.330***	0.344***
		(0.0293)	(0.0298)	(0.0309)		(0.0288)	(0.0295)	(0.0306)
$INVMODEL_{t-1}$				-1.794***				-1.836***
				(0.319)				(0.316)
$INVMODEL_{t-2}$				1.773***				1.686***
				(0.366)				(0.363)
$INVMODEL_{t-3}$				2.569***				2.709***
				(0.328)				(0.322)
MPG			71.62***	63.64***			74.56***	64.89***
			(11.81)	(12.24)			(11.65)	(12.08)
MPGxGASPRICE			-0.479***	-0.426***			-0.489***	-0.431***
			(0.0381)	(0.0386)			(0.0377)	(0.0381)
FIXED EFFECTS	Model	Model	Model	Model	Model	Model	Model	Model
TIME CONTROLS	Yes$^+$	Yes$^+$	Yes$^+$	Yes$^+$	Yes$^+$	Yes$^+$	Yes$^+$	Yes$^+$
ADDIT. CONTROLS	No	Yes^o	Yes^{o,a}	Yes^{o,a}	No	Yes^o	Yes^{o,a}	Yes^{o,a}
Observations	18,166	17,166	17,052	15,891	17,683	16,749	16,749	15,651
R-squared	0.686	0.703	0.709	0.723	0.214	0.262	0.274	0.270

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$
$^+$ indicates Time, Segment, Segment x Time controls
^{o} indicates $FLEX$, $LAUNCHED$, $PHASEDOUT$, AGE, $MSRP$, $NEWDESIGN$
^{a} indicates $GASPRICE$, $GASVOL$
Columns 5-8 use the sum of characteristics ($SIZE$, $HPWT$ and MPG) of other models of the make,
the sum of the characteristics of the models of other makes and the total number of products of other makes as IVs.
The Durbin-Wu-Hausman test of endogeneity rejects the hypothesis that PLB is exogenous (e.g., in column 8, p<0.0001)
The Angrist-Pischke multivariate F test of excluded instruments does not suggest weak instruments (e.g., in column 8, p <0.0001)

Table 11 Effect of Product Line Breadth on Mismatch Costs: Inventories. Model-level Single-equation Estimation

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
PLB	-0.368	-0.112	-0.343	0.00176	-0.0699	0.137	-0.0820	0.268
	(0.383)	(0.383)	(0.382)	(0.381)	(0.383)	(0.383)	(0.382)	(0.382)
$MPGSPAN$			-5.997***	-6.700***			-6.160***	-6.872***
			(1.394)	(1.393)			(1.374)	(1.373)
$MPGSPAN \times GASVOL$			0.128	0.111			0.149	0.132
			(0.0965)	(0.0961)			(0.0984)	(0.0979)
$DISCMODEL$				-0.00273***				-0.00268***
				(0.000280)				(0.000271)
$COMPDISC$		0.00464***	0.00295***	0.00383***		0.00435***	0.00280***	0.00369***
		(0.000600)	(0.000633)	(0.000637)		(0.000594)	(0.000625)	(0.000628)
MPG			3.969***	4.233***			4.053***	4.316***
			(0.568)	(0.577)			(0.566)	(0.575)
$MPGxGASPRICE$			-0.0216***	-0.0230***			-0.0220***	-0.0234***
			(0.00193)	(0.00198)			(0.00193)	(0.00198)
$FIXED\ EFFECTS$	Model	Model	Model	Model	Model	Model	Model	Model
$TIME\ CONTROLS$	Yes^{+}	Yes^{+}	Yes^{+}	Yes^{+}	Yes^{+}	Yes^{+}	Yes^{+}	Yes^{+}
$ADDIT.\ CONTROLS$	No	Yes^{o}	Yes^{o,a}	Yes^{o,a}	No	Yes^{o}	Yes^{o,a}	Yes^{o,a}
Observations	18,166	17,166	17,052	17,052	17,683	16,749	16,749	16,749
R-squared	0.432	0.439	0.459	0.462	0.156	0.161	0.189	0.195

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$
$^{+}$ indicates Time, Segment, Segment x Time controls
^{o} indicates $FLEX$, $LAUNCHED$, $PHASEDOUT$, AGE, $MSRP$, $NEWDESIGN$
^{a} indicates $MEDMPGMAKE$, $MEDMPGMAKExGASPRICE$, $GASPRICE$, $GASVOL$
Columns 5-8 use the sum of characteristics ($SIZE$, $HPWT$ and MPG) of other models of the make,
the sum of the characteristics of the models of other makes and the total number of products of other makes as IVs.
The Durbin-Wu-Hausman test rejects the hypothesis that PLB is exogenous (e.g., in column 8, p<0.0001)
The Angrist-Pischke multivariate F test of excluded instruments does not suggest weak instruments (e.g., in column 8, p <0.0001)

B. Demand Model and Cost Estimation

We propose a nested multinomial logit model of consumer demand (McFadden 1978, Cardell 1997). Each vehicle model is defined as a bundle of attributes. We define a nest as a combination of vehicle segment[1] and luxury level (e.g., luxury SUVs). Consumers first choose the nest in which they want to purchase (or the outside option of not buying any vehicle) and then choose the vehicle in the nest that gives them the highest utility, which is assumed to be linear in the vehicle attributes. The advantage of this model is that it avoids the problem of independence of irrelevant alternatives of conventional multinomial logit models, without adding too much computational burden. For the model estimation, we follow Berry (1994), who proposes the following transformation:

$$ln(s_{jt}) - ln(s_{0t}) = X_{jt}\beta + \sigma ln(s_{jt|g}) + \xi_{jt} \tag{3}$$

In this literature, market shares are usually defined in a slightly different way from the way we defined them in the paper (where we defined $MKTSH_{jt}$ as the total number of sales of product j in time period t divided by the total sales in the market in the same time period). Berry (1994) and related literature (e.g., for a similar model see Grigolon and Verboven 2014) often consider the “outside good” (no purchase) as one of the available “products”. In other words, shares are defined with respect to the size of the potential market, not with respect to total sales.[2] In Equation 3, s_{jt} denotes the market share of good j at time t defined in this manner (sales of good j at time t divided by total size of the market at time t) ; s_{0t} is the market share of the outside good at time t (fraction of potential customers who do not buy at time t); X_{jt} are the product characteristics, which include the price p_{jt} and the number of products in the line of the make that manufactures product j, PLB_{jt}, among other attributes; $s_{jt|g}$ is the market share of product j as a fraction of the total group share; and ξ_{jt} is a shock unobserved to the econometrician. We include segment-time controls and product fixed effects in our model.

We estimate this model with annual data at the model level. For the price p_{jt}, we subtract from the vehicle's manufacturer suggested retail price ($MSRP$) the average discount offered by the manufacturer during the year. While this is not the average price that the customers enjoy, since not all manufacturer discounts are passed to customers (Busse et al. 2006) and dealers can offer additional discounts, we believe it is a reasonable proxy. The product characteristics that we consider include vehicle size variables, a proxy for acceleration given by horsepower/weight, and the miles per gallon ($SIZE$, $HPWT$, MPG).

[1] Segments include SUV, pickup, van, compact, midsize, fullsize, sport, special.

[2] As a reference level, we set the monthly market size to 2 million. If total sales are below the market size, the difference between the market size and the total sales are purchasers of the outside good.

We account for price endogeneity using instrumental variables along the lines of the variables described in Berry et al. (1995), including characteristics of the other models by the same manufacturer and the characteristics of the rest of the vehicles on the market (for more details, see Berry et al. 1995). Note that the log of the within group share, $ln(s_{jt|g})$, is also endogenous, and therefore additional instruments, such as the number of vehicles in the nest and the characteristics of other models in the nest, are necessary. In summary, we use variation in the market shares, choice set (introduction and removal of models), vehicle attributes, and automaker product line breadth to identify the coefficients of the demand model. Table 12 shows the estimates of the demand model.

We use the estimates of this demand model to estimate the production costs, so that we can study how product line breadth affects production costs. This approach is based on an equilibrium-pricing model that arises from the demand model introduced above. Following Berry (1994), we assume that observed list prices are the result of an interior, pure strategy Nash equilibrium in prices. For the nested logit demand model, it is possible to characterize the equilibrium markup for a given model. It can be shown (see Berry 1994) that, for model j and time t, this markup is:

$$MARKUP_{jt} = \frac{(1-\sigma)/\alpha}{(1-\sigma s_{jt/g} - (1-\sigma)s_{jt})}, \tag{4}$$

where α is the price elasticity and the other parameters have been introduced before. Under this model, a model's markup depends on the market share, the within-nest market share, the substitution parameter σ, and the price elasticity α. The market share and the within-nest market share are observed in our data. The other two parameters, α and σ, are estimated using the demand model described above. Under this model, each vehicle has a different markup that takes into account market-power considerations that would be ignored if one simply proxies production costs with list prices. Using the estimated markup and the observed price for a particular model, we can recover the production costs as $COST_{jt} = PRICE_{jt} - MARKUP_{jt}$. Instead of using list prices, we consider the variable $PRICE_{jt}$, calculated as the list price minus the average discounts during the model year, since that is the price information that enters the demand model used to calculate the markup.

Table 12 Demand Estimation

	(1)
$PRICE$	-4.95e-06
	(1.05e-05)
$L\_MKSHINSEGY$	0.886***
	(0.0920)
$SIZE$	1.54e-05
	(1.11e-05)
$HPWT$	0.107
	(0.782)
MPG	0.00302
	(0.00237)
PLB	-0.00582**
	(0.00253)
Model Fixed Effects	Yes
Other controls	Segment-time
Observations	1,403
R-squared	0.998

References

Berry, S., J. Levinsohn, A. Pakes. 1995. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society* 841–890.

Berry, S.T. 1994. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics* 242–262.

Busse, M., J. Silva-Risso, F. Zettelmeyer. 2006. 1,000 cash back: The pass-through of auto manufacturer promotions. *The American Economic Review* **96**(4) 1253–1270.

Cardell, N. 1997. Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. *Econometric Theory* **13**(02) 185–213.

Grigolon, Laura, Frank Verboven. 2014. Nested logit or random coefficients logit? a comparison of alternative discrete choice models of product differentiation. *Review of Economics and Statistics* **96**(5) 916–935.

McFadden, D. 1978. *Modelling the choice of residential location*. Institute of Transportation Studies, University of California.

Rossi, Peter E. 2014. Invited paper - Even the rich can make themselves poor: A critical examination of iv methods in marketing applications. *Marketing Science* **33**(5) 655–672.