# Firms and Collective Reputation: a Study of the Volkswagen Emissions Scandal*

Rüdiger Bachmann
University of Notre Dame
CEPR, CESifo, and ifo

Gabriel Ehrlich
University of Michigan

Ying Fan
University of Michigan
CEPR and NBER

Dimitrije Ruzic
INSEAD

Benjamin Leard
University of Tennessee

July 17, 2022

**Abstract**

This paper uses the 2015 Volkswagen (VW) emissions scandal as a natural experiment to provide evidence that collective reputation externalities are economically significant. Using a combination of difference-in-differences and demand estimation approaches, we document a spillover effect from the scandal to the non-VW German auto manufacturers. The spillover amounts to an average drop of $2,057 in consumer valuations of these manufacturers' vehicles and to a 34.6% reduction in their annual sales. We substantiate our interpretation that the estimates reflect a reputation spillover using data on internet search behavior and direct measures of consumer sentiment from Twitter.

**JEL Codes**: D12, L14, L62.
**Keywords**: automobiles, collective reputation, demand estimation, difference-in-differences, Google trends, reputation externalities, Twitter sentiment, Volkswagen emissions scandal.

*We would like to thank Networked Insights for providing data, as well as seminar and conference participants for helpful discussions. All errors are our own. E-mail contact: rbachman@nd.edu, gehrlich@umich.edu, yingfan@umich.edu, dimitrije.ruzic@insead.edu, or bleard@utk.edu. This version of the paper supersedes its first version titled "Firms and Collective Reputation: the Volkswagen Emissions Scandal as a Case Study" and published as CEPR-DP 12504 and CESifo-WP 6805 in December 2017.

"Collective reputations play an important role in economics and the social sciences. Countries, ethnic, racial or religious groups are known to be hard-working, honest, corrupt, hospitable or belligerent. Some firms enjoy substantial rents from their reputations for producing high-quality goods."

— Jean Tirole, 1996

# 1 Introduction

Reputation plays an important role in mitigating problems that arise from incomplete information. While firms have individual reputations, groups of firms may also share collective reputations. These collective reputations can come about through formal associations such as franchises or through informal associations such as country-of-origin labels, for instance French wine, Swiss watches, or German engineering. With collective reputations come externalities: the action of one franchised store may affect the reputation of the whole chain; similarly, an industrial scandal implicating one firm might spill over to other firms in the industry. While such relationships among firms are common and frequently in the public eye, measuring the externalities arising from collective reputations remains challenging. In this paper, we identify and quantify the economic significance of collective reputation externalities. We do so in the context of a major industrial scandal in the United States, the 2015 Volkswagen (VW) emissions scandal.

In addition to being widespread, collective reputations merit attention both because economists have often theorized that they play an important role in how the economy functions, and because they raise questions about market discipline for firm misbehavior. First, as Tirole (1996) notes, "Collective reputations play an important role in economics and the social sciences," and there is a well-developed theoretical literature on the topic.[1] Nonetheless, the literature that empirically documents collective reputations among firms and measures the resulting externalities remains thin. Second, the externalities arising from collective reputation imply that firms may not fully bear the costs or reap the benefits of decisions that affect the collective reputation of their group. In that sense, there may be a public goods problem in which firms underinvest in behaviors that build and sustain collective reputations. Conversely, firms may be insufficiently averse to engaging in behaviors that, if discovered, would cause harmful reputational spillovers. The collective reputation externalities we document in this paper thus provide a counterpoint to the common argument that consumer responses to firm misbehavior effectively discipline firms.

[1] In addition to Tirole (1996), see for example Levin (2009) and Neeman, Öry and Yu (2019).

A challenge in identifying and measuring collective reputation spillovers is that products are potentially substitutes. Consider a stylized example of a market with three firms. Two firms share a collective reputation (e.g., VW and BMW are associated with "German Engineering") while a third does not (e.g., Ford). Suppose a negative shock harms the reputation of the first firm and that the shock spills over to harm the reputation of the second firm as well (e.g., misbehavior at VW tarnishes the reputation of "German Engineering" and hence BMW). At the same time, substitution away from the first firm may drive demand toward the second firm (BMW) as well as the third firm (Ford). These substitution effects generate two challenges to measuring the reputation spillover correctly. First, no firm in the market is completely "untreated" by the shock. Second, the observed change in sales for the second firm (BMW) is the combination of two potential effects: the reputational spillover effect of interest and the substitution effect.

We overcome these challenges by combining difference-in-differences and structural estimation approaches and by providing direct evidence on the scandal's reputational consequences in the United States. First, we use a difference-in-differences approach to show that the VW emissions scandal reduced the U.S. vehicle sales of the non-VW German auto manufacturers—BMW, Mercedes-Benz, and Smart—relative to their non-German counterparts.[2] We consider this result to be initial evidence that there was a country-specific spillover from the VW scandal to the other German auto manufacturers. Second, we estimate a model of vehicle demand with flexible substitution patterns to quantify the spillover and the substitution effects separately. Third, we provide supplementary evidence that this spillover effect is reputational in nature.

To relate the two approaches, we provide a framework to conceptualize the difference between the difference-in-differences estimates and the true spillover effect of an industrial scandal. We clarify how the two are related and provide conditions on the patterns of substitution under which difference-in-differences results constitute evidence for the existence of a spillover effect. In our setting, the estimated patterns of demand confirm the validity of these conditions. In other settings, researchers may lack the rich data needed to estimate flexible substitution patterns but have supplementary information about these patterns. Our framework establishes how and when difference-in-differences estimates can indicate the existence of a spillover effect in those settings.

Two features of the VW emissions scandal make it an attractive natural experiment to study spillovers from collective reputation; first is the scandal's sudden revelation and intense media prominence as one of the largest industrial scandals in recent history. To summarize, on September 18, 2015, the U.S. Environmental Protection Agency (EPA) served

[2]Opel, a German auto manufacturer formerly owned by General Motors, does not sell in the United States.

a Notice of Violation to the VW Group alleging that approximately 500,000 VW and Audi diesel-engine vehicles sold between 2009 and 2015 in the United States contained a defeat device allowing these vehicles to appear to comply with emissions regulations in the test box, while having higher on-road emissions.[3] For the general public, the scandal was a clear surprise in September 2015, and it immediately generated extensive media coverage. Moreover, the scandal occurred in an important industry in the United States, of which German vehicles constitute a large share.[4]

Second, the automotive sector features salient makes, well-developed tools for studying vehicle demand, and rich available data. In addition to individual automotive makes being salient to consumers, the German auto manufacturers featured the notion of "German engineering" prominently in their U.S. advertising, creating a natural reputational group. To capture consumer substitution patterns across vehicles—and thereby separate substitution from spillovers—we leverage the tools that the industrial organization literature has developed for estimating vehicle demand. Our estimation procedure uses detailed data not only on vehicle sales and characteristics, but also surveys that describe consumers' demographics and their second-choice vehicles.[5]

We find an economically significant reputational spillover from the scandal, amounting to an average drop of $2,057 in consumer valuations of the non-VW German manufacturers' vehicles and a 34.6% reduction in those manufacturers' annual U.S. unit sales. The scandal's total effect is smaller than its spillover effect because the spillover and substitution effects move in opposite directions. While the reputational spillover harmed the non-VW German auto manufacturers, substitution away from VW benefited them. We estimate that substitution away from VW led to an 11.0% increase in the other German auto manufacturers' unit sales, implying that the scandal's net total effect (a 23.5% reduction in sales) was smaller than the reputational spillover on its own.

The finding that the spillover effect coexists with a countervailing substitution effect away from VW—and that, therefore, the spillover effect is larger in absolute value than the scandal's combined effects—is unlikely to be a coincidence. Firms that are associated closely enough to have a collective reputation (so that the spillover effect exists) are also likely to produce close substitutes (which determines the substitution effect). Our results

[3]The Volkswagen Group consists of Volkswagen proper plus Audi and Porsche.

[4]In 2014, German auto manufacturers accounted for 8.1 percent of all U.S. light vehicle sales, making Germany the second-largest source for foreign-branded vehicles.

[5]Specifically, we combine product-level data from WardsAuto with household-level survey data from a major market research company. The addition of household-level data—which includes information on demographics, actual purchase decisions, and second-choice purchase intentions—aids our estimation of substitution patterns by providing direct information about: (1) which types of households purchase which types of vehicles, and (2) which pairs of vehicles are close substitutes for particular households.

show the economic importance of these spillovers and indicate that the spillovers can be systematically undermeasured when substitution patterns are not taken into account.

We supplement our estimates of the scandal's effects with Twitter data to document changes in sentiment showing harm to the non-VW German automakers' collective reputation. These directly-observable changes in sentiment support our argument that the estimated German-specific spillover was a reputational spillover. Moreover, we show patterns of internet search behavior that cast doubt on an alternative interpretation of the spillover effect that works through information rather than reputation. Finally, we provide evidence that consumers at the time would have had no technical or economic reason to believe that the other German auto manufacturers were implicated in the scandal.

Our work contributes to the literature on reputation by studying collective reputation and by quantifying the reputation spillovers on the economic outcomes of firms. The focus on group reputation distinguishes our paper from the literature on individual and platform reputations, including Cabral and Hortaçsu (2010), Li (2010), Mayzlin, Dover and Chevalier (2014), Nosko and Tadelis (2015), Fan, Ju and Xiao (2016), Luca (2016), and Li, Tadelis and Zhou (2020).[6] Our paper also relates to work studying the reputational effects of industrial scandals, such as Jonsson, Greve and Fujiwara-Greve (2009) for the Swedish finance industry, Freedman, Kearney and Lederman (2012) for the U.S. toy industry, Barrage, Chyn and Hastings (2020) for the British Petroleum oil spill, and Bai, Gazze and Wang (2021) for the Chinese dairy industry.

Our work also incorporates reputational spillovers in a large literature on demand estimation in the auto industry (e.g., Berry, Levinsohn and Pakes, 1995; Petrin, 2002; Berry, Levinsohn and Pakes, 2004; Train and Winston, 2007; Li, 2019; Springel, 2020; Xing, Leard and Li, 2021). We use a similar combination of automotive and survey data to estimate our model as Grieco, Murry and Yurukoglu (2021), who study the evolution of market power in the U.S. auto industry. Our discrete-choice model allows consumers to value certain vehicle characteristics—such as a VW nameplate, a diesel engine, or a non-VW German origin—differently before and after the VW scandal. The coefficients on the latter two characteristics allow the scandal to have potential spillovers on vehicles with diesel engines and on the German reputational group.

From the perspective of individual firms, collective reputations can arise both exogenously, because firms share salient characteristics, and endogenously, through various mar-

[6]There is also a finance literature that studies how a variety of corporate events adversely affect firm values and interprets such effects as reputational losses; see, for example, Fiordelisi, Soana and Schwizer (2014) for a summary of this literature. We have conducted a similar analysis using stock prices in a previous version of the paper (Bachmann, Ehrlich, Fan and Ruzic, 2019). The analysis shows that the VW emissions scandal reduced the cumulative abnormal stock returns of the non-VW German auto manufacturers.

keting approaches. Newburry (2012) emphasizes the importance of "country of origin" effects for multinational firms, while geographical designations are pervasive for agricultural products (Winfree and McCluskey, 2005). Fishman, Finkelstein, Simhon and Yacouel (2018) and Neeman, Öry and Yu (2019) rationalize endogenous reputational group formation by showing that group reputations reduce the likelihood that consumers will develop extreme beliefs (positive or negative) about firm quality, thus providing a form of reputational risk pooling. Castriota and Delmastra (2014) argue that, if there are economies of scale in advertising and promotion, groups of firms may be better positioned to incur those expenses than individual producers.

Additionally, there is a small but growing literature that studies the consequences of the Volkswagen emissions scandal. Strittmatter and Lechner (2020), Che, Katayama and Lee (2018), and Ater and Yoseph (2020) study the scandal's effect on Volkswagen vehicles in the used car market. Griffin and Lont (2018) and Barth, Eckert, Gatzert and Scholz (2019) investigate the financial market effects for VW and other large automakers. Alexander and Schwandt (2019) use the scandal as a natural experiment to study the health effects of vehicle exhaust. Relatedly, Alé-Chilet, Chen, Li and Reynaert (2021) study a subsequent collusion case in which three German automakers coordinated on the size of diesel exhaust fluid tanks.

The remainder of the paper is organized as follows: Section 2 provides a more detailed explanation and timeline of the VW emissions scandal and describes the scandal's effect on VW. Section 3 provides difference-in-differences estimates that indicate the existence of a German-specific spillover effect from the scandal. Section 4 presents a model of vehicle demand and quantifies the spillover effect. Section 5 provides support for our interpretation of the spillover effect as a reputational spillover, and discusses alternative interpretations. A final Section 6 concludes.

## 2 The VW Emissions Scandal as a Natural Experiment

In this section, we describe the timeline of the VW emissions scandal in more detail and argue that it provides a good setting to study the spillovers arising from collective reputation. Using data from print publications, the stock market, and social media, we show that the scandal was largely unanticipated. We then provide evidence substantiating the claim that German auto manufacturers share a group identity.

## 2.1 Timeline of the Scandal

In May 2014, West Virginia University's Center for Alternative Fuels Engines and Emissions found discrepancies between high on-road emissions by VW diesel vehicles and earlier test results. The EPA and the California Air Resources Board (CARB) permitted a voluntary recall of VW diesel vehicles in December 2014. In May 2015, CARB conducted new tests, and again the on-road emissions failed to match the test-box results for VW diesel vehicles. In July 2015, the agencies informed VW about these tests and threatened not to certify the 2016 diesel vehicles. On September 3, 2015, VW admitted to the EPA and CARB that it had used a defeat device in its software, which regulated emissions and produced fake test results in the test box (see Breitinger (2018) for a more complete timeline). The scandal entered its public phase on September 18, 2015, when the EPA served a Notice of Violation to the Volkswagen Group.

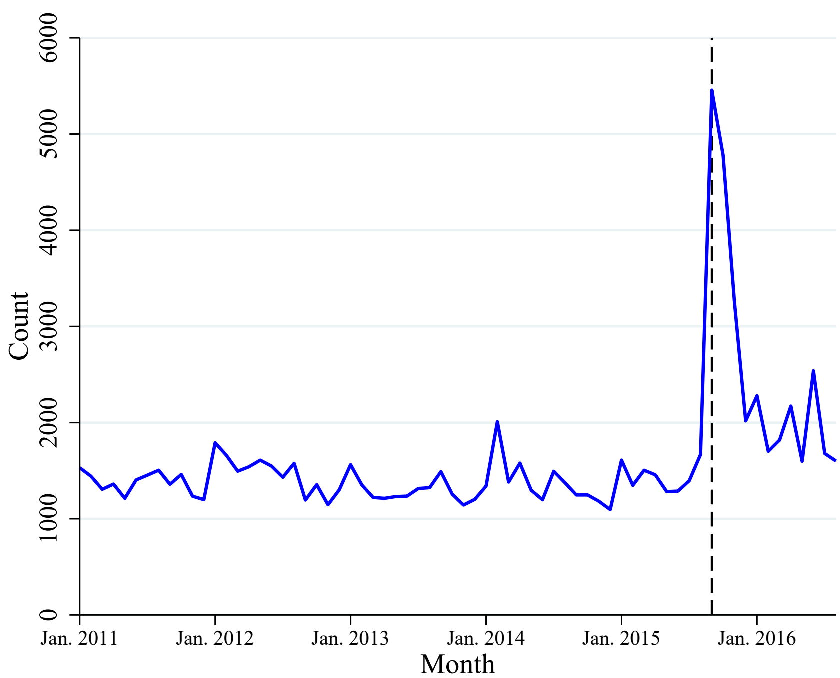
Volkswagen's culpability quickly became a matter of public knowledge: on September 20, two days after the start of the scandal, Volkswagen admitted publicly to the deception and issued an apology. VW Chief Executive Officer Martin Winterkorn resigned three days later, on September 23. He was eventually charged with fraud in the United States in May 2018. On September 28, German authorities opened a fraud investigation of the former CEO, and in October they authorized a police raid on the VW headquarters. The U.S. Congress called the VW U.S. CEO, Michael Horn, to testify on October 8, 2015, and he formally resigned his post in early March 2016. In anticipation of the fines and settlements associated with the scandal, VW set aside more than $18 billion in fiscal year 2015. The scandal's legal resolution in the United States began in April 2016. On July 26, 2016, VW and a U.S. court agreed on a civil settlement totaling $15 billion.

Major news outlets across many countries covered the scandal and its aftermath. On September 19, the morning after the scandal, the front page of the New York Times read: "U.S. Orders Major VW Recall Over Emissions Test Trickery." The Wall Street Journal used a more accusatory tone: "Volkswagen Faked EPA Exhaust Test, U.S. Alleges." Spiegel Online and Zeit Online, the online platforms of two major German newspapers, frequently reported about the scandal, which also quickly spilled over into popular culture. For example, on October 13, 2015, Paramount Pictures and Leonardo DiCaprio's production company announced that they had secured the rights to shoot a film about the scandal, and on September 22, 2016, VW was awarded the satirical Ig Noble Prize in chemistry (Improbable Research, 2016).

## 2.2 The Scandal Surprised the General Public

Monthly print media mentions of "Volkswagen" more than tripled in September 2015, suggesting that the scandal came as a complete surprise to the general public. We quantify the media prominence of the scandal using data from the Newsbank news aggregator on print media mentions of "Volkswagen" in the United States. The database covers roughly 5,000 U.S. newspapers, newswires, journals, and magazines. Figure 1 shows that mentions of "Volkswagen" spiked from a pre-scandal monthly average of 1,500 to 5,500 in September 2015. This sudden increase suggests that the scandal caught the media and public by surprise.

Figure 1: Monthly Print Media Mentions of "Volkswagen" in the United States




Note: The dashed line shows the month of the Volkswagen emissions scandal, September 2015. The data come from the Newsbank news aggregator, which covers roughly 5,000 U.S. newspapers, newswires, journals, and magazines. The time period covered is January 2011 to August 2016.

Along with the adverse attention in the media, VW's stock price declined precipitously following the EPA's announcement; the visually evident discontinuity on September 18 in Figure 2 suggests that the scandal came as a surprise to market participants. Volkswagen's end-of-day stock price fell by 33 percent in the two trading days following the scandal. The stock price subsequently recovered some of its losses over the rest of the year, but at the end of August 2016 it remained 24 percent lower than its pre-scandal closing price.[7]

[7] To focus on the effects within the United States and to avoid currency effects from the euro-based VW listing on the Frankfurt Stock Exchange, we use the price of the VW American Depository Receipt (ADR) traded on U.S. markets. ADRs are issued by a U.S. depository bank, entitle the owner to shares in an international security and are priced and pay dividends in U.S. dollars.

Figure 2: End-of-Day Stock Price for Volkswagen Group

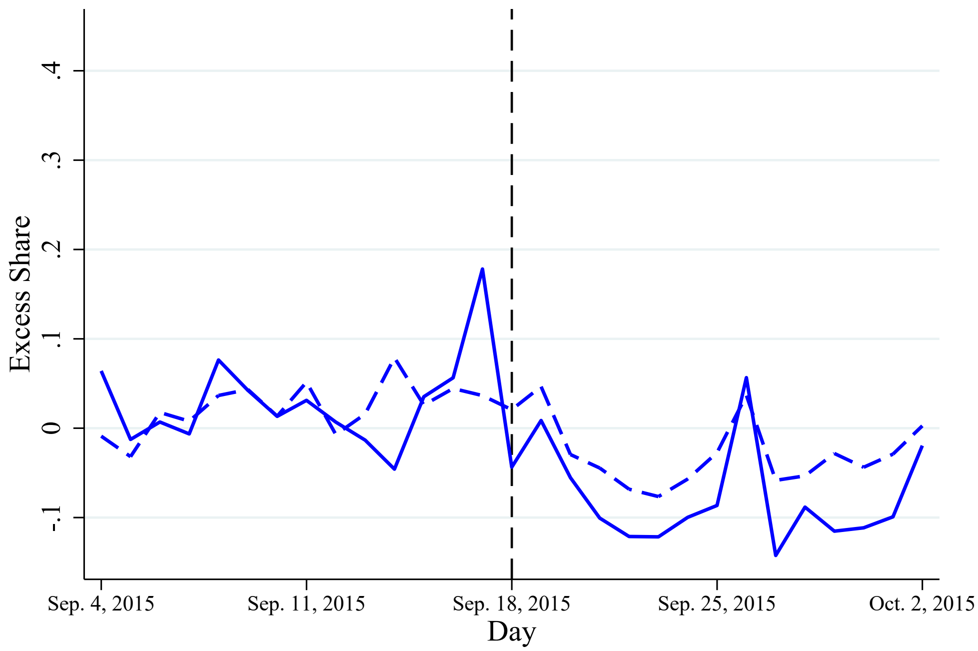

Note: The dashed line shows the date of the Volkswagen emissions scandal, dated September 18, 2015. The end-of-day price shown for Volkswagen ADR is listed on U.S. stock exchanges. The data come from the Bloomberg database. The time period covered is January 2011 to August 2016.

Furthermore, the tone of social media discussion regarding Volkswagen suddenly shifted, with positive sentiment declining and negative sentiment spiking in the aftermath of the scandal. We document this pattern with novel sentiment measures from Networked Insights.[8] We focus on sentiment data from Twitter, an online social media networking service where roughly 300 million active monthly users share short messages. The sentiment measures in our data set are calculated from a 10 percent random sample from Twitter. Networked Insights categorizes tweets as displaying positive, neutral, or negative sentiment toward the mentioned company. Posts are excluded from the analysis if they are not written in English or if the user accounts are associated with locations outside the United States. Networked Insights also constructs brand identifiers. An identifier for Volkswagen, for instance, is meant to collate mentions of "Volkswagen," "VW," "#Volkswagen," and the like. Given the size of the underlying data set, Networked Insights only retains the past 13 months of data. We requested the data in September 2016, so our time series begins on August 10, 2015, a little over a month before the scandal became public. We first create average daily sentiment shares (positive/negative/neutral) for August 2015 for each

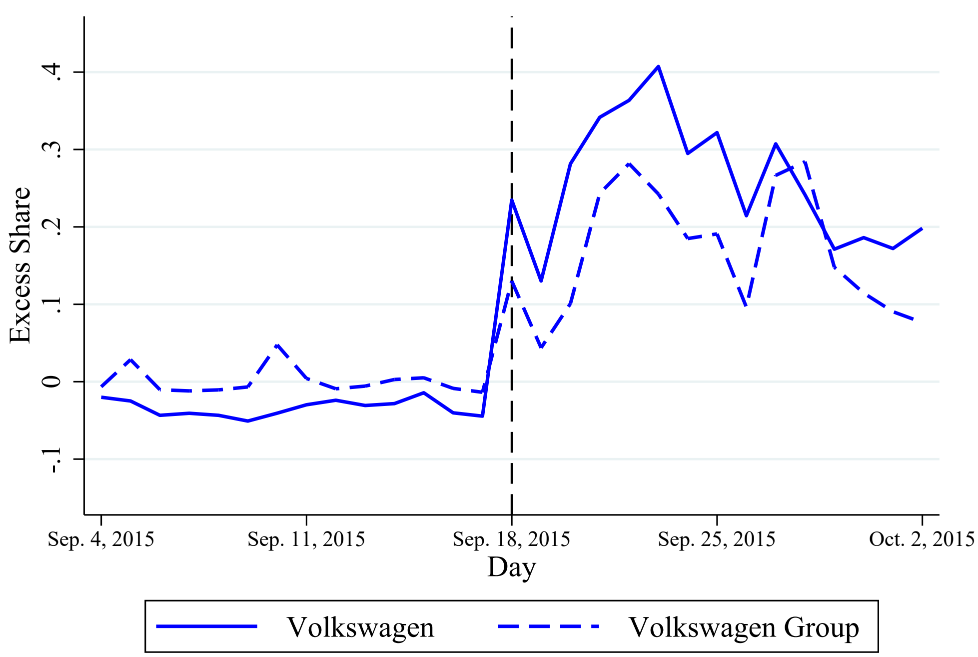
[8]Networked Insights is a data analytics company, founded in 2006, that provides a platform for real-time semantic analyses of social media posts; its primary clients are consumer-facing companies that use the platform to manage their brands.

Figure 3: Daily Twitter Sentiment Towards Volkswagen

A: Positive Sentiment in Excess of August 2015 Average




B: Negative Sentiment in Excess of August 2015 Average

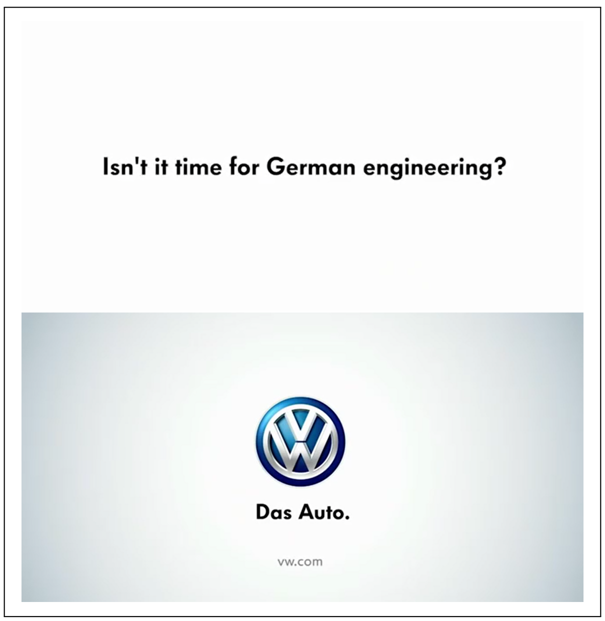



Note: The dashed vertical lines show the date of the Volkswagen emissions scandal, dated September 18, 2015. The figure shows the normalized shares of tweets expressing positive/negative Twitter sentiment towards a particular make. The denominator of these shares includes positive, negative and neutral sentiments. Sentiment shares are normalized by subtracting the average sentiment share from August 10 to August 31, 2015. We show a window of $\pm 14$ days around September 18, 2015. Volkswagen Group is defined as Volkswagen, Audi, and Porsche. The data come from Networked Insights.

vehicle make in our data to serve as a pre-scandal baseline. We then construct sentiment shares in excess of this August baseline for each day.

Figure 3 displays these sentiment metrics for VW and the VW Group two weeks before and after the scandal. Panel A shows a decrease in positive sentiment toward VW, from an average of 3 percentage points higher than its August baseline in the two weeks prior to the scandal to an average of 8 percentage points below in the two weeks following the scandal. Panel B displays a sharp increase in negative sentiment toward VW, from an average of 3 percentage points below to an average of 26 percentage points above.[9] The results for the entire Volkswagen group (which includes Audi and Porsche) are similar. Together, these two panels suggest that Volkswagen's reputation suffered in the aftermath of the September 18 EPA announcement.

## 2.3 "German Engineering" as a Group Identity

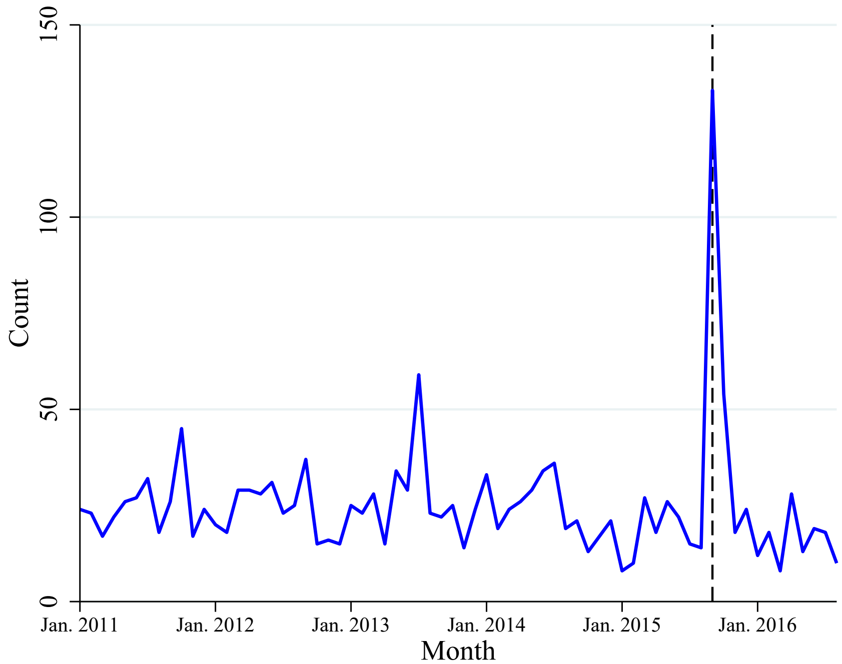



Having established that the VW scandal was a shock and that it affected VW's reputation, we now provide *prima facie* evidence that there is a collective German reputation through which the scandal may have had a spillover effect on other German automakers.

We first note that German auto manufacturing companies have historically leveraged the broader reputation of "German engineering" in their marketing. For instance, a VW commercial from 2014 states, "... Everyone knows that the best cars in the world come from Germany." The ad fades out to the question: "Isn't it time for German engineering?", and then pivots to the German phrase "Das Auto" ("The Car"), presumably in order to associate VW and "German engineering" with the idea of the archetypical car.

[9]The pre-scandal and post-scandal means are statistically different at the 1 percent significance level for both positive and negative sentiment.

Figure 4: Monthly Print Media Mentions of "German Engineering" in the United States




Note: The dashed line shows the month of the Volkswagen emissions scandal, September 2015. The data come from the Newsbank news aggregator, which covers roughly 5,000 U.S. newspapers, newswires, journals, and magazines. The time period covered is January 2011 to August 2016.

It is, therefore, not surprising that following the scandal, media attention to "German engineering" spiked, with 130 print articles mentioning the term in September 2015, a five-fold increase over the preceding months. We illustrate this increase in Figure 4 using data from the Newsbank aggregator. A recurring theme in this news coverage was the notion that the scandal might tarnish the broader reputation of German manufacturing firms. As part of this coverage of the scandal, Reuters published an article on September 22, 2015, titled "VW scandal threatens 'Made in Germany' image" (Chambers, 2015). A day later, Reuters doubled down with an article titled "Volkswagen could pose bigger threat to German economy than Greek crisis" (Nienaber, 2015), which included the claim: "The broader concern for the German government is that other car makers such as Mercedes-Benz and BMW could suffer fallout from the Volkswagen disaster."[10]

[10] See also Bruckner (2015), Werz (2016), and Remsky (2017).

# 3 Difference-in-Differences Evidence on the Spillover Effect

In this section, we present difference-in-differences evidence that the VW emissions scandal indeed had a spillover effect on the other German auto manufacturers (BMW, Mercedes-Benz, and Smart). We show that the scandal substantially reduced the U.S. sales growth of the other German automakers relative to their non-German counterparts. We then interpret our difference-in-differences estimates through the lens of a general demand system to clarify how they relate to spillovers.

## 3.1 Regression Results

To study the effects of the VW emissions scandal, we obtain data on U.S. light vehicle sales from WardsAuto. WardsAuto receives sales data from all auto manufacturers in the United States. It is thus in principle a complete count of light vehicle sales in the United States.[11] An individual observation in the data contains identifiers for the vehicle make (e.g., Honda or Volkswagen), the vehicle model (e.g., Civic or Jetta), and the vehicle powertype (e.g., gas or diesel). The estimation uses data on 37 makes, listed in Appendix Table A.1, and 357 distinct models. We identify six makes as German in origin: Audi, BMW, Mercedes-Benz, Porsche, Smart, and Volkswagen.[12] The sample period starts in January 2010 and ends in August 2016, one year after the scandal. Focusing on the period immediately after the scandal helps limit potential contamination from subsequent shocks in the automotive industry, such as the unrelated scandals we discuss in Section 5.

We use a difference-in-differences regression specification to show how the scandal affected German auto manufacturers relative to the non-German auto manufacturers. We begin the analysis by constructing a total sales measure for each make, so that an observation is a make-month (e.g., Honda in January 2016). Following a standard difference-in-differences regression specification (e.g., Angrist and Krueger, 1999), we estimate the following regression:

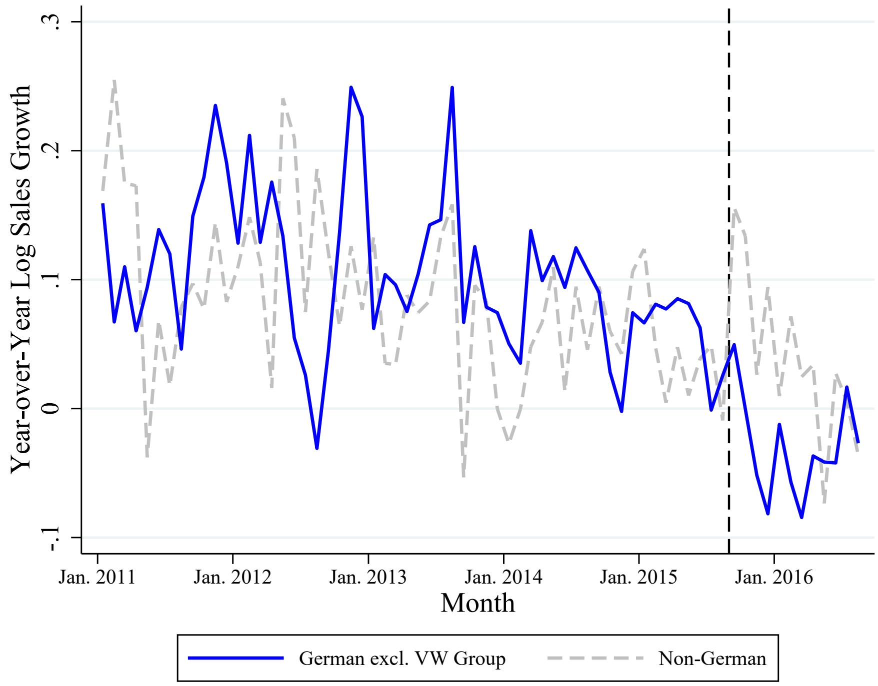
$$y_{kt} = \eta_k + \gamma_t + \rho T_{kt} + \varepsilon_{kt}, \tag{1}$$

where $y_{kt} = \ln \mathit{Sales}_{kt} - \ln \mathit{Sales}_{kt-12}$ is the 12-month log sales growth rate of vehicle make $k$ at time $t$. $\eta_k$ is a make-specific fixed effect, capturing potential make-level heterogeneity in growth rates. $\gamma_t$ is a fixed effect for each month in the sample, capturing potential

[11] The official U.S. vehicle sales statistics in national accounting data are based on the same data we use.

[12] Mini, the present-day incarnation of a line manufactured by the British Motor Corporation and its successors between 1959 and 2000, is currently owned by BMW. Given its historical association with Britain, we classify Mini as not of German origin. We consider alternative classifications in Appendix B.1 and show that the results are not sensitive to this choice.

Figure 5: Differences in U.S. Light Vehicle Sales Growth




Note: The dashed vertical line shows the month of the Volkswagen emissions scandal, September 2015. The data come from WardsAuto. Volkswagen Group is defined as Volkswagen, Audi, and Porsche.

seasonality in vehicle sales growth and the potential impacts of time-varying fuel prices. $T_{kt}$ is an indicator taking value one for the German auto manufacturers during and after the scandal month, and zero otherwise. The coefficient of interest, $\rho$, captures the scandal's differential impact on non-VW German auto manufacturers relative to non-German auto manufacturers.

We exclude the Volkswagen Group from the sample to focus the analysis on the economic consequences of reputation for German automakers not directly implicated by the scandal. We weight this regression by the square root of sales volumes to dampen the impact of highly volatile sales growth rates of small sales levels. Figure 5 shows that the pre-scandal trends in sales growth for the non-VW German and non-German auto manufacturers were comparable.[13] Figure 5 also shows that the sales growth of the non-VW German manufacturers turned negative following the scandal, offering a first indication that there was an adverse spillover effect from the scandal.

The estimation results in Table 1 show that the scandal reduced the sales growth rates of the non-VW German automakers by 9.2 percentage points relative to their non-German counterparts (Column 1). Given the scandal's origins in the diesel market, a natural concern is that this estimated effect could be driven by substitution away from diesel vehicles,

[13] Appendix B.2 adds some additional evidence for this assessment.

Table 1: Difference-in-Differences Estimates
German vs. Non-German Auto Manufacturers, Excl. VW Group

| Dependent Variable | 12-month Log Sales Growth | |
|---|---|---|
| Power Type | Baseline | non-Diesel |
| | (1) | (2) |
| German × Post-Scandal | -0.092*** | -0.084** |
| | (0.030) | (0.033) |
| Time Fixed Effects | Yes | Yes |
| Make Fixed Effects | Yes | Yes |
| $R^2$ | 0.303 | 0.300 |
| N | 2150 | 2150 |

Note: The unit of observation is a make-month (e.g., the log growth of all BMW sales from January 2014 to January 2015). The time period covered is January 2011 to August 2016. Standard errors clustered at the make level in parentheses. VW Group (VW, Audi, and Porsche) is excluded from all regressions. The VW emissions scandal is dated September 18, 2015. Sales are measured in units sold. All regressions are weighted by the square root of sales volumes. The data come from WardsAuto. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

especially given that German and non-German auto manufacturers differ in their exposure to the diesel market. Column (2) assuages such concerns by repeating the difference-in-differences regression for non-diesel vehicle sales only. We find that the scandal reduced German automakers' sales growth of non-diesel vehicles by 8.4 percentage points relative to that of the non-German auto manufacturers. These results suggest that substitution away from diesel cannot, on its own, explain the scandal's relative effects on the other German automakers: an additional channel must have been at play.[14]

## 3.2 Interpreting the Difference-in-Differences Estimate

In this section, we propose a framework to clarify how difference-in-differences (DID) estimates such as those in the previous section relate to the true spillover effect of an industrial scandal. Although a DID approach may misestimate the true spillover effect because of substitution, we provide a condition under which a DID estimate indicates that a spillover exists. We argue that this condition is likely to apply in many settings in which researchers use scandals to study spillovers from collective reputations.

[14]In Appendix B.3 we show that the relative decline in the non-VW German automakers' sales holds for alternative econometric specifications such as an unweighted regression, calculating sales growth rates relative to the average of sales in consecutive periods, using a shorter pre-scandal period, or using log sales levels rather than log differences as the dependent variable.

We consider a market with three groups of firms producing potentially substitutable products: first, the "source" firm, which is directly implicated in scandalous behavior; second, a group of "treated" firms, which are not directly implicated but which share a collective reputation with the source firm; and third, a group of "comparison" firms, which are not implicated in the scandal and do not share a collective reputation with the source firm. Let $\theta_{source}$ and $\theta_{spillover}$ denote, respectively, the deterioration of the source and treated firms' reputations following the scandal.

In what follows, we define an outcome of interest for firm $k$ as $y_k(\theta_{source}, \theta_{spillover})$. Implicitly, the function $y_k$ may depend on each individual firm's and its competitors' product prices and characteristics as well as consumer demographics and incomes. We suppress this dependence in the notation in this section for expositional simplicity. We assume the following properties of this outcome function:

$$\frac{\partial y_k}{\partial \theta_{source}} \geq 0 \text{ for all } k \in \{\text{treated, comparison}\} \tag{2a}$$

$$\frac{\partial y_k}{\partial \theta_{spillover}} \leq 0 \text{ for all } k \in \{\text{treated}\}, \quad \frac{\partial y_k}{\partial \theta_{spillover}} \geq 0 \text{ for all } k \in \{\text{comparison}\}. \tag{2b}$$

For instance, $y_k$ may be the demand for firm $k$'s products (or a monotonic transformation of demand). In that case, these properties are implications of any demand model in which products are substitutes, so that a firm's demand increases in its own reputation and decreases in its competitors' reputations.

**What a DID Estimate Captures**

We define the true spillover effect on the outcome of interest as:

$$\Omega_{spillover} \triangleq \underset{k \in \text{treated}}{\text{mean}} [y_k(\theta_{source}, \theta_{spillover}) - y_k(\theta_{source}, 0)].^{15} \tag{3}$$

By contrast, the differences-in-differences estimator $\Omega_{DID}$ compares the change in treated firms' outcomes before and after the scandal to the change in the comparison firms' outcomes. An empirical DID estimate may include controls for time-varying factors other than the scandal. Therefore, in this stylized framework, we can define a DID estimate as shown

[15] If the spillover were defined in the absence of the "source" shock, i.e., setting $\theta_{source} = 0$ in equation (3), all of the arguments in this section would still apply. We need only to replace $y_k(\theta_{source}, 0)$ by $y_k(0, \theta_{spillover})$ in equation (5) and swap the labels for the $\Omega$ terms in each row.

in equation (4), which can be decomposed into the four terms in equation (5):

$$\Omega_{DID} \triangleq \underset{k\in\text{treated}}{\text{mean}}[y_k(\theta_{source},\theta_{spill.}) - y_k(0,0)] - \underset{k\in\text{comparison}}{\text{mean}}[y_k(\theta_{source},\theta_{spill.}) - y_k(0,0)] \tag{4}$$

$$= \Bigg( \underbrace{\underset{k\in\text{treated}}{\text{mean}} \Big[y_k(\theta_{source},\theta_{spill.}) - y_k(\theta_{source},0)\Big]}_{\text{spillover effect on treatment group } \triangleq \Omega_{spillover}} + \underbrace{\underset{k\in\text{treated}}{\text{mean}} \Big[y_k(\theta_{source},0) - y_k(0,0)\Big]}_{\text{substitution effect on treatment group } \triangleq \Omega_2} \Bigg) \tag{5}$$
$$- \Bigg( \underbrace{\underset{k\in\text{comparison}}{\text{mean}} \Big[y_k(\theta_{source},\theta_{spill.}) - y_k(\theta_{source},0)\Big]}_{\text{substitution effect on comparison group due to } \theta_{spill.} \triangleq \Omega_3} + \underbrace{\underset{k\in\text{comparison}}{\text{mean}} \Big[y_k(\theta_{source},0) - y_k(0,0)\Big]}_{\text{substitution effect on comparison group } \triangleq \Omega_4} \Bigg).$$

The first term, $\Omega_{spillover}$, is the spillover effect of interest, while the second and fourth terms, $\Omega_2$ and $\Omega_4$, arise as consumers substitute away from the source firm; by inequality (2a), they are both positive. The third term, $\Omega_3$, captures how the comparison group's outcome changes as a result of the spillover's harm to the treatment firms' reputations; both $\Omega_{spillover}$ and $\Omega_3$ would be zero in the absence of a spillover.

This decomposition illustrates three challenges to inferring $\Omega_{spillover}$ from $\Omega_{DID}$. First, the comparison group may be treated by the shock (i.e., $\Omega_3, \Omega_4 \neq 0$), so ostensibly "untreated" firms may not be appropriate controls for the treatment group. Second, substitution away from the source firm toward the treatment group will bias $\Omega_{DID}$ away from finding a spillover effect (i.e., $\Omega_2 \geq 0$ by inequality 2a). Third, following the shock to the source firm, even if there were no true spillover effect ($\theta_{\text{spillover}} = 0$ and thus $\Omega_{spillover} = \Omega_3 = 0$), $\Omega_{DID}$ may still be non-zero because of differential substitution patterns across the treatment and comparison groups.

### Relationship between DID Estimates and Spillovers

Despite these challenges, we provide an intuitive condition under which the $\Omega_{DID}$ estimator guarantees the *existence* of a spillover even if it does not provide a quantitatively accurate estimate of the spillover.

**Assumption 1** *The shock to the source firm, $\theta_{source}$, induces a larger substitution effect toward the treatment group than toward the comparison group (i.e., $\Omega_2 > \Omega_4$).*

**Proposition 1** *Under Assumption 1, if the DID estimate is negative (i.e., $\Omega_{DID} < 0$), then the spillover effect exists (i.e., $\theta_{spillover} > 0$ and $\Omega_{spillover} < 0$).*

**Proof.** Assumption 1 (i.e., $\Omega_2 > \Omega_4$) implies that if $\Omega_{DID}(= \Omega_{spillover} + \Omega_2 - \Omega_3 - \Omega_4) < 0$, then $\Omega_{spillover} - \Omega_3 < 0$. Note that inequalities (2b) imply that $sign(\Omega_{spillover}) \neq sign(\Omega_3)$. Therefore $\Omega_{spillover} < 0$ (and $\Omega_3 > 0$). By (2b) we also have that $\theta_{spillover} < 0$. ∎

We believe that Assumption 1 is natural in settings with collective reputations, because products close enough to share a collective reputation are likely to be closer substitutes than products from other groups.[16] Although the U.S. auto market features rich data, researchers attempting to estimate firm reputation spillovers in other settings may be less fortunate. Proposition (1) shows that, in those cases, difference-in-differences estimation can indicate the existence of a spillover effect under an assumption about substitution patterns, especially when researchers have supplementary knowledge regarding likely substitution patterns.

In the next section, we estimate a model of vehicle demand, and we use it to quantify the VW emissions scandal's spillover effects separately from its substitution effects. Those estimates confirm Assumption 1 in the context of the VW emissions scandal, i.e., substitution away from VW Group sales following the scandal was disproportionately toward the other German auto manufacturers. Therefore, in line with our Proposition 1, we argue that our difference-in-differences estimates indicate the existence of a spillover effect from the scandal.

# 4 Quantifying the Spillover Effect: Demand Estimation and Decomposition

In this section, we estimate a model of vehicle demand that features flexible substitution patterns as a means of disentangling the spillover and substitution effects of the VW emissions scandal. We use this model to decompose the scandal's impact on the U.S. sales of the non-VW German automakers into three potential forces: a substitution effect, a spillover effect, and a diesel effect. We allow for a diesel effect in the model because of diesel technology's role at the center of the scandal.

## 4.1 Demand Model

We assume that U.S. households' vehicle demand is described by a discrete-choice model. The indirect utility that household $i$ derives from purchasing a vehicle of make-

[16]Under a stronger set of conditions, we can also discuss the *quantitative* relationship between $\Omega_{DID}$ and the true spillover effect, $\Omega_{spillover}$. Equation (5) shows that if the comparison group were not affected by the scandal (i.e., $\Omega_3 = \Omega_4 = 0$), then $\Omega_{DID} = \Omega_{spillover} + \Omega_2$, where $\Omega_2 \geq 0$. By continuity, if the firms in the comparison group are weak substitutes to the source and treatment firms (so that $\Omega_3, \Omega_4 \approx 0$), then the DID estimate $\Omega_{DID}$ will be a lower bound for the scandal's spillover harm to the treatment group (i.e., $\Omega_{DID} > \Omega_{spillover}$, noting that a negative value of $\Omega_{spillover}$ represents harm).

model-power type $j$ (hereafter, product $j$) in year $t$ is given by:

$$u_{ijt} = x_{jt}\beta_i + \xi_{jt} + \varepsilon_{ijt}, \tag{6}$$

where the vector $x_{jt}$ includes the price of product $j$ in year $t$, key observable attributes of each product such as miles per gallon and vehicle weight, and fixed effects for country of manufacturer and power type. Of particular importance for our purposes, $x_{jt}$ also contains three dummy interaction terms, $\mathbf{1}^+$(*Scandal Make*$_j$ $\times$ *Post-Scandal*$_t$), $\mathbf{1}^+$(*Other German*$_j$ $\times$ *Post-Scandal*$_t$) and $\mathbf{1}^+$(*Diesel*$_j$ $\times$ *Post-Scandal*$_t$). These interaction terms allow for potential changes in consumer valuations of makes in the VW Group, for spillovers to other German makes, and for potential changes in consumer valuations of diesel engines.[17] The term $\xi_{jt}$ captures product characteristics that are not observable to the econometrician but are known to households and auto manufacturers. Finally, $\varepsilon_{ijt}$ is an idiosyncratic taste shock. We normalize the utility of household $i$'s outside option of not purchasing a vehicle during year $t$ to $u_{i0t} = \varepsilon_{i0t}$.

To flexibly capture substitution patterns, we allow households to have heterogeneous tastes with respect to product prices and characteristics. Specifically, we assume that the coefficient vector $\beta_i$ can vary with household demographics $d_i$ such as income and household size, so that $\beta_i = \beta + \phi d_i$.

We also allow households' idiosyncratic tastes to be correlated across products. Specifically, we assume that idiosyncratic tastes $\varepsilon_{ijt}$ follow a Generalized Extreme Value distribution that allows for correlation in $\varepsilon_{ijt}$ for products of the same country origin.[18] Let $\lambda \in [0, 1)$ be the within-country correlation parameter. A $\lambda$ equal to zero represents the case of no differential correlation. Positive estimates of $\lambda$ would suggest that idiosyncratic tastes are more closely correlated for products that share a common country origin.

To derive choice probabilities for households and market shares for products, we first define the mean utility over all households that purchase product $j$ in year $t$ as:

$$\delta_{jt} = x_{jt}\beta + \xi_{jt}. \tag{7}$$

We denote by $\mathcal{J}_t$ the set of products available in year $t$, by $\mathcal{J}_{gt}$ the set of products with country origin group $g$, and by $(\boldsymbol{\delta}_t, \boldsymbol{x}_t)$ the collection of $(\delta_{jt}, x_{jt})$ for all $j \in \mathcal{J}_t$. The prob-

[17] The product-level dummy variables *Scandal Makes*$_j$, *Other German*$_j$, and *Diesel*$_j$ take the value $1$ if, respectively, product $j$ belongs to the Volkswagen group, belongs to other German makes, or uses diesel fuel, and take the value $0$ otherwise. The time-varying dummy variable *Post-Scandal*$_t$ takes the value $1$ if year $t$ is after the scandal and $0$ before the scandal.

[18] In other words, $\varepsilon_{ijt}$ follows a nested Logit structure where a group (nest) consists of products with the same country origin: Germany, Japan, Korea, United States, and All Other country origins. The outside option is its own group.

ability that household $i$ chooses product $j$ of group $g$ is $\Pr(u_{ijt} \geq u_{ij't}, \forall j' \in \mathcal{J}_t)$, which is given by:

$$Pr_{jt}(\boldsymbol{\delta}_t, \boldsymbol{x}_t, d_i; \phi, \lambda) = \frac{exp\{(\delta_{jt} + x_{jt}\phi d_i)/(1-\lambda)\}}{I_{igt}^{\lambda}[1 + \sum_{g'} I_{ig't}^{(1-\lambda)}]}, \tag{8}$$
$$\text{where } I_{igt} = \sum_{j' \in \mathcal{J}_{gt}} exp\{(\delta_{j't} + x_{j't}\phi d_i)/(1-\lambda)\}.$$

Aggregating this choice probability using the distribution of demographics $d_i$ across households, denoted by $G_t(d_i)$, gives the market share of each product as follows:

$$s_{jt}(\boldsymbol{\delta}_t, \boldsymbol{x}_t; \phi, \lambda) = \int Pr_{jt}(\boldsymbol{\delta}_t, \boldsymbol{x}_t, d_i; \phi, \lambda) dG_t(d_i). \tag{9}$$

In Section 4.2, we describe how we map this model market share to the data as a means of estimating the key model parameters. We also use this market share function for quantifying the effects of the scandal in Section 4.3.

## 4.2 Data and Estimation

### 4.2.1 Combining Product and Household Data

To estimate our model of vehicle demand, we combine sales and characteristic data at the product-year level from WardsAuto and survey data at the household level from a major market research company. As noted, a product in our analysis is a make-model-power type combination (e.g., a Honda Civic with a gasoline engine). We define market size by the number of households in the United States, and we use the all-items Consumer Price Index for all urban consumers published by the Bureau of Labor Statistics to express prices in constant 2015 dollars.

The WardsAuto data provide information on unit sales, prices (manufacturers' suggested retail prices), and key product attributes for all products between 2010 and 2016. While the sales volume data is at the monthly frequency, the product characteristics data (including prices) is at the annual frequency. Consequently, following much of the broader literature on estimating vehicle demand, we estimate the model at the annual frequency. Because the VW emissions scandal became public in September 2015, we define all years prior to 2015 as the pre-scandal period and 2016 as the post-scandal period; we exclude 2015 from our estimation sample. Furthermore, the vehicle characteristics data are reported at the product-trim level, while the sales volume data are at the product level. We thus aggregate the characteristics data by taking the minimum across trims within each product. As a robustness analysis, we also consider the median across trims. Appendix C.1 provides summary statistics of prices and product attributes.

Our household-level survey data provides information on households who have purchased new vehicles. For each household in the dataset, we observe household demographics (i.e., household income, size, and the age of the respondent), the vehicle purchased and other vehicles considered (e.g., vehicle "most seriously considered"). Consistent with the WardsAuto data, we use the survey data for the years 2010 to 2016.

### 4.2.2 Estimation Procedure

We carry out the estimation in two steps, as in Goolsbee and Petrin (2004) and Train and Winston (2007).[19] In the first step, we leverage the household-level data and use maximum likelihood to estimate the parameters that capture heterogeneity in consumer tastes: $\phi$, which governs how consumer tastes vary with demographics, and $\lambda$, the within-country correlation of idiosyncratic taste shocks. We also obtain the mean utility $\delta_{jt}$ for each product. In the second step, we use an instrumental-variables approach to regress mean utility $\delta_{jt}$ on product characteristics $x_{jt}$ to obtain an estimate of $\beta$, the parameters that capture the mean tastes of consumers.

To conduct the first step, we leverage the fact that households in the surveys report both the vehicle they purchased as well as the vehicle that the household "most seriously considered" to construct the likelihood function with which we estimate parameters $\phi$ and $\lambda$. We interpret those two survey responses as the household's first- and second-most preferred product choices (see Berry, Levinsohn and Pakes (2004) for a similar interpretation of similar survey data from General Motors).[20] Let $j_i^1$ and $j_i^2$ represent household $i$'s first and second choices, respectively. The joint probability of household $i$ choosing $j_i^1$ as its first choice and $j_i^2$ as its second choice is the probability that $j_i^2$ is a better choice than any $j' \notin \{j_i^1, j_i^2\}$, minus the probability that $j_i^2$ is a better choice than any $j' \neq j_i^2$. In other

[19] Grieco, Murry, Pinkse and Sagl (2021) suggest combining the product-level market share data and the household-level survey data into one likelihood function for estimation. As pointed out in their paper, the efficiency gain from doing so is modest when the size of the household-level survey data is small compared to the market size, which is the case in our paper. Moreover, as opposed to Goolsbee and Petrin (2004) and Murry and Zhou (2020), our market share data are not constructed from the household-level survey data. They come from the WardsAuto data, which report the actual total vehicle sales in the United States. Therefore, we think there is merit in putting more weight on the market share data by matching the observed market shares perfectly, following Berry, Levinsohn and Pakes (1995).

[20] The survey data provides information on the third and fourth choices as well. Due to a large number of missing observations for these choices, we use only the first- and second-choice data.

words, the probability is $\Pr(u_{ij_i^2t} \geq u_{ij't}$ for $\forall j' \in \mathcal{J}_t \backslash j_i^1) - \Pr(u_{ij_i^2t} \geq u_{ij't}$ for $\forall j' \in \mathcal{J}_t)$,[21] where the latter is given by equation (8) and the former can be analogously computed from the choice set excluding $j_i^1$. With a slight abuse of notation, we rewrite the choice probability in (8) as $Pr_j(\boldsymbol{\delta}_t, \boldsymbol{x}_t, d_i; \mathcal{J}_t; \phi, \lambda)$ to make its dependence on the choice set $\mathcal{J}_t$ explicit. Then the likelihood of observing household $i$'s first two choices as $j_i^1$ and $j_i^2$ is the following difference:

$$Pr_{j_i^2}(\boldsymbol{\delta}_t, \boldsymbol{x}_t, d_i; \mathcal{J}_t \backslash j_i^1; \phi, \lambda) - Pr_{j_i^2}(\boldsymbol{\delta}_t, \boldsymbol{x}_t, d_i; \mathcal{J}_t; \phi, \lambda). \tag{10}$$

Instead of maximizing the likelihood over the entire parameter space of $(\phi, \lambda, \boldsymbol{\delta})$ directly, we invert out the vector of mean utilities $\boldsymbol{\delta}(\phi, \lambda)$ and search over the parameter space $(\phi, \lambda)$. Specifically, we invert out the vector of mean utilities $\boldsymbol{\delta}$ by matching the observed market shares in the data and the market shares predicted by the demand model (i.e., equation 9), in the spirit of Berry (1994) and Berry, Levinsohn and Pakes (1995).[22] Let $\boldsymbol{\delta}_t(\phi, \lambda)$ be the solution to $s_{jt}(\boldsymbol{\delta}_t, \boldsymbol{x}_t; \phi, \lambda) = s_{jt}$ for all $j \in \mathcal{J}_t$, where the right-hand side is the observed market share. Plugging $\boldsymbol{\delta}_t(\phi, \lambda)$ into (10) yields the likelihood function for estimation:

$$\begin{aligned} L(\phi, \lambda) = \sum_t \sum_{i \in \mathcal{I}_t} log\Big( & Pr_{j_i^2}(\boldsymbol{\delta}_t(\phi, \lambda), \boldsymbol{x}_t, d_i; \mathcal{J}_t \backslash j_i^1; \phi, \lambda) \\ & - Pr_{j_i^2}(\boldsymbol{\delta}_t(\phi, \lambda), \boldsymbol{x}_t, d_i; \mathcal{J}_t; \phi, \lambda)\Big), \end{aligned} \tag{11}$$

where $\mathcal{I}_t$ represents the households in a specific year's survey. We randomly sample 10,000 households from each year of the survey to form the likelihood function (11). This step also produces estimates of the mean utilities $\delta_{jt}$.

To conduct the second step, we regress the estimated mean utility of each product $\delta_{jt}$ on its price and characteristics using equation (7) and thereby estimate $\beta$, the parameters capturing the average taste. Because product prices may be correlated with their unobservable demand shocks (i.e., the error term $\xi_{jt}$ in equation 7), ordinary least squares estimation could lead to biased estimates. Consequently, we use an instrumental-variables approach. Specifically, we construct instrumental variables based on the characteristics of close competing products (Gandhi and Houde, 2019). In the oligopolistic market for

[21] This is because the probability of household $i$ choosing $j_i^1$ and $j_i^2$ as its first and second choices is $\Pr(u_{ij_i^1t} \geq u_{ij_i^2t} \geq u_{ij't}, \forall j' \in \mathcal{J}_t \backslash j_i^1)$. This probability is equivalent to $\Pr(u_{ij_i^1t} \geq u_{ij_i^2t}$ and $u_{ij_i^2t} \geq u_{ij't}, \forall j' \in \mathcal{J}_t \backslash j_i^1)$, which can be written as the difference between $\Pr(u_{ij_i^2t} \geq u_{ij't}, \forall j' \in \mathcal{J}_t \backslash j_i^1)$ and $\Pr(u_{ij_i^1t} < u_{ij_i^2t}$ and $u_{ij_i^2t} \geq u_{ij't}, \forall j' \in \mathcal{J}_t \backslash j_i^1)$. Note that the second term equals $\Pr(u_{ij_i^2t} \geq u_{ij't}, \forall j' \in \mathcal{J}_t)$.

[22] We compute the expectation in equation (9) by drawing a sample of household characteristics $d_i$ from the Current Population Survey (CPS) using the household CPS weights.

automobiles, the price of product $j$ depends not only on its own product characteristics, but also on the attributes of the other products available in the market, especially close substitutes as measured by the distance in attributes. The correlation between the characteristics of other products and the potentially endogenous price makes those products' characteristics relevant instruments. Under the timing assumption that automakers decide the characteristics of their products before the realization of the demand shocks, the characteristics of other products are uncorrelated with product $j$'s demand shock. Appendix C.2 presents the first-stage regressions and details the instrumental variables we use.

#### 4.2.3 Estimation Results

Table 2 reports the estimation results. The estimates show that consumer utility decreases with a vehicle's price and increases with a vehicle's size, its fuel efficiency, and its horsepower relative to size. We also find that price sensitivity decreases both with household income and with household age. Larger and older households tend to prefer larger vehicles. These estimates are robust to different ways of aggregating characteristics across a product's trims: for our baseline results in Column (1), we use the minimum characteristic, and in Column (2) we use the median characteristic. The results are robust across the two aggregation methods, so we focus on the baseline in what follows.

The estimated coefficient on *VW Group* $\times$ *Post-Scandal* is negative (–0.486), implying that the scandal reduced consumer valuations of VW vehicles. This coefficient equates to an average decline of $3,057 in consumer valuations of VW Group vehicles.[23]

The estimated coefficient for *Other German* $\times$ *Post-Scandal* is also negative (–0.327), implying that the scandal reduced consumer valuations of the other German manufacturers' vehicles by an average of $2,057. This estimated change in valuations reflects the scandal's spillover effect on the other German automakers.

We find little change in consumers' tastes for diesel vehicles following the VW scandal. The estimated coefficients are small and statistically insignificant. The finding that the scandal changed consumer valuations of German vehicles more noticeably than the consumer valuation of diesel is consistent with our results in Table 1, where our difference-in-differences estimates are quantitatively similar whether we consider all vehicles or focus on non-diesel sales. Moreover, the finding of a relatively small impact on diesel tastes combined with the documented reduction in valuations for VW and other German manu-

[23] We compute this decline in consumer valuations based on the change in consumer utility from purchasing a VW product post scandal and the average price coefficient. The former is the coefficient of *VW Group* $\times$ *Post-Scandal*, i.e., –0.486; the latter is the price coefficient averaged across demographics, which is 1.590. Since our price variable is denominated in $10,000, the implied decline in consumer valuations of VW products is –0.486 $\times$ $10,000/1.590 = $3,057.

Table 2: Demand Estimation Results—U.S. Light Vehicle Sales

| | Trim Aggregation: Min | Trim Aggregation: Median |
|---|---|---|
| | (1) | (2) |
| Price (\$10,000) | -2.439*** | -1.932*** |
| | (0.104) | (0.062) |
| Price × Income (\$1m) | 3.807*** | 3.531*** |
| | (0.060) | (0.070) |
| Price × Income$^2$ | -1.382*** | -1.729*** |
| | (0.092) | (0.141) |
| Price × Age | 1.096*** | 0.917*** |
| | (0.017) | (0.017) |
| *ln* Weight | 5.319*** | 4.721*** |
| | (0.444) | (0.312) |
| *ln* Weight × 1(Household Size≥3) | 0.176*** | 0.171*** |
| | (0.010) | (0.010) |
| *ln* Weight × Income | -2.055*** | -2.543*** |
| | (0.143) | (0.144) |
| *ln* Weight × Income$^2$ | -1.451*** | -0.544*** |
| | (0.199) | (0.191) |
| *ln* Weight × Age | 1.317*** | 1.270*** |
| | (0.030) | (0.043) |
| Horsepower/Weight (HP/1000lb) | 0.035*** | 0.025*** |
| | (0.006) | (0.004) |
| *ln* Miles-Per-Gallon | 1.521*** | 1.610*** |
| | (0.164) | (0.165) |
| VW Group × Post-Scandal | -0.486*** | -0.391** |
| | (0.159) | (0.159) |
| Other German × Post-Scandal | -0.327* | -0.413** |
| | (0.175) | (0.165) |
| Diesel × Post-Scandal | 0.071 | 0.145 |
| | (0.241) | (0.239) |
| Nest parameter | 0.395*** | 0.396*** |
| | (0.002) | (0.002) |
| Country Fixed Effects | yes | yes |
| Power-Type Fixed Effects | yes | yes |

Note: Column (1) uses the minimum characteristic across trims as the product characteristic, and Column (2) uses the median. A product is defined as a make-model-power type combination (e.g., a Honda Civic with a gasoline engine). The time period covered is 2010 to 2016, with 2015 omitted as the scandal year. Each calendar year is considered a distinct market. Country Fixed Effects comprises indicators for German, Japanese, Korean, U.S., and All Other country origins. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

facturers suggests that the spillover from the scandal (and hence any associated change in reputation) belongs to makes and countries and not to products *per se*. Our results are thus consistent with Newburry (2012), who emphasizes that country of origin is a key element of reputation for multinational firms.

The estimated nest parameter $\lambda$ is 0.395, indicating that, conditional on vehicle characteristics, consumers viewed vehicles within the same country grouping as closer substitutes than vehicles across country groupings. Overall, our estimated demand model implies that the decline in consumer valuations of the VW Group led to a 12% increase in the non-VW German auto manufacturers' sales, while Japanese, Korean, U.S., and "Other National Origin" makes saw unit sales increase by only 0.6% to 0.7%. These results indicate that German vehicles are closer substitutes than vehicles of other national origins, consistent with Assumption 1 from Section 3.2, which we used to interpret the difference-in-differences results as indicating the existence of a German-specific spillover from the scandal.

On the whole, the estimation results suggest three channels through which the VW scandal affected the sales of the non-VW German auto manufacturers. First, we estimate a significant decline in consumer valuations of VW vehicles. In a market where differentiated vehicles are (imperfect) substitutes, this decline in valuations drove sales away from VW toward vehicles produced by other manufacturers. Furthermore, this force drove sales disproportionately toward the other German auto manufacturers, because their vehicles are closer substitutes to VW than are non-German vehicles. We refer to this increase in sales as the "substitution effect." Second, consumer valuations of the other German auto manufacturers also declined, reducing their sales. We refer to this decline in sales as the "spillover effect." Third, we allow but find little support for the idea that the scandal changed consumers' taste for diesel vehicles. In the next section, we quantify the scandal's overall effects on the non-VW German manufacturers' vehicle sales and revenues, as well as the relative contributions of each of these channels.[24]

## 4.3 Quantifying the Effects

We now quantify how the scandal and its individual channels shifted vehicle demand for the non-VW German auto manufacturers. To do so, we consider the U.S. light vehicle market from 2014, the year prior to the scandal, and simulate how vehicle sales would have changed had one or more of the scandal's effects been present. We hold the set of products and their characteristics—including prices—fixed in this exercise. Our exercise is thus akin to asking how far the demand curve shifts in a simple supply-demand diagram;

[24]In Appendix D, we show that the patterns documented throughout this section are robust to different specifications of time, country, and make fixed effects.

it is distinct from asking how the scandal affected equilibrium outcomes. We focus on this demand quantification as a transparent way to isolate the VW emissions scandal's individual channels. Nonetheless, we also show that our decomposition is robust to an alternative quantification of the scandal's effects that considers equilibrium outcomes.

Table 3 summarizes the design of the three main simulations we use to quantify the scandal's effects. As the scandal took place in 2015, the 2014 data reflect none of those effects. In Simulation 1, we turn on all three channels by setting all three post-scandal interaction dummies to one and then recompute vehicle sales. Comparing the simulated sales and revenues from Simulation 1 to those in the data gives us the overall effect of the scandal. In Simulation 2, we turn on only two channels of the VW scandal: the diesel channel and the spillover channel. Specifically, we set to one the post-scandal interaction dummies for changes in the consumer valuations of the non-VW German auto manufacturers and for changes in the consumer valuation of diesel; we leave as zero the remaining interaction dummy for the change in consumer valuations for VW Group vehicles after the scandal. The difference between Simulations 1 and 2 is therefore only in whether substitution away from the VW Group is present; comparing outcomes for these two simulations thus quantifies the substitution effect. Simulation 3 allows for only one channel: the spillover driven by the estimated change in consumers' tastes for the non-VW German auto manufacturers. The comparison between Simulations 2 and 3 isolates the diesel effect that is present in Simulation 2 but not in Simulation 3. Finally, because Simulation 3 includes only the country-specific spillover channel, the difference between Simulation 3 and the (pre-scandal) data quantifies the spillover effect.

Table 3: Simulation Designs

| | Simluation 1 | Simulation 2 | Simulation 3 | Data |
|---|---|---|---|---|
| Substitution Effect | yes | no | no | no |
| Diesel Effect | yes | yes | no | no |
| Spillover Effect | yes | yes | yes | no |

Table 4 reports the overall effects of the scandal on non-VW German auto manufacturers and a decomposition of the overall effects into the three channels. Together, the scandal's three channels reduce the sales of non-VW German auto manufacturers by 165,965 units worth $7.7 billion, roughly 23.5% of their 2014 sales. This decline is driven by the spillover effect, which lowers sales for the non-VW German automakers by 244,607 vehicles and reduces their revenues by $11.2 billion. This effect is partially offset by the

Table 4: Scandal's Impact on Other German Manufacturers

| | Comparison | Vehicle Sales | Revenue ($ billion) |
|---|---|---|---|
| Overall Scandal | Simulation 1 − Data | –165,965 | –7.74 |
| | | –23.50% | –23.66% |
| Substitution Effect | Simulation 1 − Simulation 2 | 77,342 | 3.44 |
| | | 10.95% | 10.52% |
| Diesel Effect | Simulation 2 − Simulation 3 | 1,300 | 0.06 |
| | | 0.18% | 0.17% |
| Spillover Effect | Simulation 3 − Data | –244,607 | –11.24 |
| | | –34.63% | –34.35% |

Note: Simulation designs are defined in Table 3. Revenue effects are expressed in 2015 dollars.

substitution effect, which increases sales for the non-VW German auto manufacturers by 77,342 units, worth $3.4 billion, as consumers switch away from Volkswagen. The changes in unit sales and revenue ascribed to the changing consumer valuation for diesel are much smaller than the substitution and spillover effects. Because the spillover and substitution effects move in opposite directions, and the diesel effect is quantitatively negligible, the scandal's total net effect is substantially smaller than the spillover effect individually.

Our finding that the spillover effect is significant and larger than the overall effect of the VW scandal is robust to alternative decompositions of the scandal's impact on vehicle demand. The three simulations we consider in this section are not the only way to isolate the scandal's effects. For instance, in Table 4, we quantify the spillover effect by comparing Simulation 3 (where only the spillover channel is active) with the data (where no channel is active). However, the comparison of any two simulations that differ only in whether the spillover channel is active would also be a valid quantification. In principle, the non-linearity of the demand model means that the order in which each channel is turned on or off could affect the decomposition results. We find, however, that the possible ranges for the individual effect sizes over all possible decompositions, including the baseline decomposition presented in Table 4, are quantitatively narrow. Therefore, we conclude that the order in which consider the individual channels is practically unimportant for isolating their effects. We detail all possible simulations and comparisons and report confidence intervals in Appendix C.3.

While our primary interest is in quantifying the scandal's impact on demand, our findings are also robust to considering price adjustments. We now repeat our previous simulations allowing firms to adjust prices and report the results in Table 5.[25] Doing so requires additional assumptions about the supply side of the model. Specifically, we model firms as setting prices in a Bertrand competition and back out the (constant) marginal cost for each product based on the first-order conditions with respect to prices. We then recompute the equilibrium prices while turning on some or all of the scandal's channels, using the backed-out marginal costs. Appendix C.4 provides full model and estimation details. In this alternative quantification, the equilibrium change in vehicles sold comes both from a shift in the demand curve and from a movement along the demand curve. By contrast, the baseline quantification in this section isolated the scandal's channels by restricting attention to shifts in consumer demand. Our main findings are robust across these two quantifications.

Table 5: Scandal's Impact on Other German Manufacturers with Price Adjustments

| | Comparison | Vehicle Sales | Revenue ($ billion) |
|---|---|---|---|
| Overall Scandal | Simulation 1 − Data | –156,655 | –7.30 |
| | | –22.18% | –22.32% |
| Substitution Effect | Simulation 1 − Simulation 2 | 81,820 | 3.66 |
| | | 11.58% | 11.17% |
| Diesel Effect | Simulation 2 − Simulation 3 | 1,022 | 0.04 |
| | | 0.14% | 0.11% |
| Spillover Effect | Simulation 3 − Data | –239,498 | –10.99 |
| | | –33.91% | –33.60% |

Note: Simulation designs are defined in Table 3. Revenue effects are expressed in 2015 dollars.

[25]Another margin of adjustment could be firms' advertising strategies. Unfortunately, data limitations prevent us from studying firms' advertising strategies. We believe, however, that our focus on the short-run aftermath of the scandal limits the scope for such changes to affect our results.

# 5 Interpretation of the Spillover Effect

In this section, we argue that the most plausible interpretation of the spillover effect documented and quantified in the previous two sections is that it arises from a collective reputation shared by all German automakers selling vehicles in the United States. As mentioned in Section 2, German auto manufacturers used the notion of "German engineering" in their advertising, providing *prima facie* evidence that they share a collective reputation. Here, we begin by presenting more direct evidence that the VW emissions scandal led to a deterioration in public perceptions of the other German automakers sharing a collective reputation. We then discuss an alternative information-based interpretation and argue that it is less persuasive. Finally, we describe subsequent scandals involving the German automakers that occurred after the end of our analysis period, and thus conclude that these subsequent scandals cannot explain our documented spillover effect.

## 5.1 Social Media Evidence on Reputation Spillover

We use the Twitter sentiment data described in Section 2.2 to show that perceptions of the non-VW German automakers suffered in the aftermath of the VW emissions scandal. Specifically, we estimate the difference-in-differences specification in equation (1) with Twitter sentiment data for the non-VW German automakers. The unit of observation is a make-day.[26] Column (1) of Table 6 shows a statistically significant decline of 3.5 percentage points in positive sentiment toward non-VW German auto manufacturers as a result of the scandal (we discuss the results for negative Twitter sentiment in Column (2) later in this section). To put this number in perspective, the share of tweets expressing positive sentiment toward those companies averaged 12.3 percent in August 2015. We consider this deterioration in positive social-media sentiment towards the other German automakers as suggesting that the sales spillover effect we document arose from a collective reputation.

## 5.2 Discussion of an Alternative Interpretation

An alternative interpretation of the spillover effect to the non-VW German auto manufacturers is that, following the scandal, consumers suspected these other German auto manufacturers of having engaged in cheating behavior similar to VW's. Such a suspicion

[26]The Networked Insights database does not include identifiers for some makes in the WardsAuto data. Table A.2 in Appendix A lists the makes with Twitter sentiment data. The estimation sample is a window of $\pm 14$ days around the scandal eruption date of September 18, 2015, and the outcome variable is the share of tweets expressing positive/negative Twitter sentiment towards a particular make. The denominator of these shares includes positive, negative and neutral sentiments.

Table 6: Difference-in-Differences Estimates—Twitter Sentiment
German vs. Non-German Auto Manufacturers, Excl. VW Group

| Dependent Variable | Positive Sentiment | Negative Sentiment |
|---|---|---|
| | (1) | (2) |
| German × Post-Scandal | -0.035 | 0.002 |
| | (0.006) | (0.006) |
| Time Fixed Effects | Yes | Yes |
| Make Fixed Effects | Yes | Yes |
| $R^2$ | 0.348 | 0.268 |
| N | 840 | 840 |

Note: The unit of observation is a make-day. Sentiment shares are normalized by subtracting the average sentiment share from August 10 to August 31, 2015. The denominator of these shares includes positive, negative and neutral sentiments. The estimation period comprises 14 days before and after scandal date of September 18, 2015. Volkswagen Group is defined as Volkswagen, Audi, and Porsche. All regressions include make and time fixed effects, and are weighted by tweet volume. The data come from Networked Insights.

could have arisen from a shared German identity, in which case this alternative channel is ultimately connected to group reputation. If, however, such a suspicion arose from independent information available to consumers around the time of the scandal, this alternative interpretation would indeed be distinct from our notion of collective reputation.

We present three arguments against this alternative interpretation. Our first argument provides evidence suggesting that any such "shared suspicion" was not substantiated by information available to consumers around the time of the VW scandal. Our second argument is that patterns of social media sentiment are indeed inconsistent with consumers possessing information that implicated the other German auto manufacturers in malfeasance similar to VW's. Our third argument goes further and presents online search patterns suggesting that consumers did not possess such suspicions in the first place.

First, there was no concurrent notion of malfeasance by the non-VW German automakers. The West Virginia University study that ultimately led to the discovery of the VW scandal focused on three diesel vehicles: a VW Passat, a VW Jetta, and a BMW X5. The VW vehicles failed the test, whereas the BMW vehicle passed. No Mercedes-Benz vehicles were tested (see Thompson, Carder, Besch, Thiruvengadam and Kappanna, 2014). In addition, Mercedes-Benz and BMW had little technical reason to resort to cheating devices, as their diesel vehicles tended to be larger than VW's. As a result, they used exclusively the Selective Catalytic Reduction (SCR) system as their $NO_x$-control system, which adds urea and water to the exhaust flows. This system is more effective and more reliable in reducing $NO_x$ emissions but requires more space for the urea and water containers. VW, with on average smaller vehicles in its fleet, developed a different system, a nitrogen oxide trap, which the VW engineers could never get to operate with the same efficiency as SCR systems.[27] This difficulty was what led to VW's deception in the first place: VW would not have been able to comply with U.S. regulations without cheating (Zycher, 2017). Finally, media at the time wrote that non-VW German automakers were not implicated in the scandal. On September 22, 2015, for instance, CNN wrote (Petroff, 2015): "But before you start worrying about the complete collapse of the German auto industry, it's worth repeating that—at least for now—the scandal is limited to Volkswagen. Other German automakers such as Daimler, which owns Mercedes-Benz, and BMW have said they're not affected." Even in the fall of 2017, Zycher wrote (Zycher, 2017): "Note that VW is the only manufacturer accused of explicitly installing such systems to defeat $NO_x$ emissions control systems. [...] In short, it is perverse to fail to distinguish between the behavior of VW and that of the rest of the industry."

Second, Twitter sentiment toward the other German automakers displayed a very different pattern after the scandal than did Twitter sentiment toward VW. As mentioned in Section 2.2, negative sentiment toward VW increased sharply following the scandal, while positive sentiment declined moderately. This evidence suggests that suspicion of malfeasance manifests mainly as an increase in negative sentiment towards the wrongdoer. Therefore, if consumers indeed had information that other German automakers were cheating too, we should expect a similar pattern for their social-media sentiment. However, as shown in Column (2) of Table 6, we find no meaningful change in negative Twitter sentiment following the scandal for the other German automakers. By contrast, positive sentiment toward them declined, as shown in Column (1). This decline in positive sentiment without a corresponding increase in negative sentiment suggests that perceptions of non-VW German automakers suffered after the scandal for reasons other than consumers possessing information about malfeasance.

Third, although online searches for Volkswagen spiked following the scandal, consumers showed no heightened inquisitiveness toward the other German auto manufacturers. Figure 6 contains four panels, each of which plots a time series of a single Google

[27] Of the three tested vehicles, only the VW Passat had a SCR system installed, but, according to Zycher (2017), the system was operating with reduced efficiency so as not to inconvenience the driver with too frequent refills of the urea.

search term ("Volkswagen", "VW", "BMW", and "Mercedes").[28] Searches for "Volkswagen" and "VW" in panels A and B increased dramatically in the aftermath of the scandal, with respective z-scores of 22 and 15 in the week of September 18, the date of the EPA announcement. By contrast, searches for the two main non-VW German makes, "BMW" in panel C and "Mercedes" in panel D, were indistinguishable from their regular fluctuations. The Google search data implies that consumers did not become more inquisitive about the other German auto manufacturers following the scandal. If consumers suspected those automakers of cheating similar to VW's, they did not display it in a similar accompanying search for information.

Taken together, these three arguments suggest that the scandal's effects on non-VW German automakers were unlikely to be driven by information. Rather, the scandal must have tarnished the reputations of the other German auto manufacturers through their association with Volkswagen, consistent with the notion of a collective reputation.

## 5.3 Subsequent Scandals in the Auto Industry

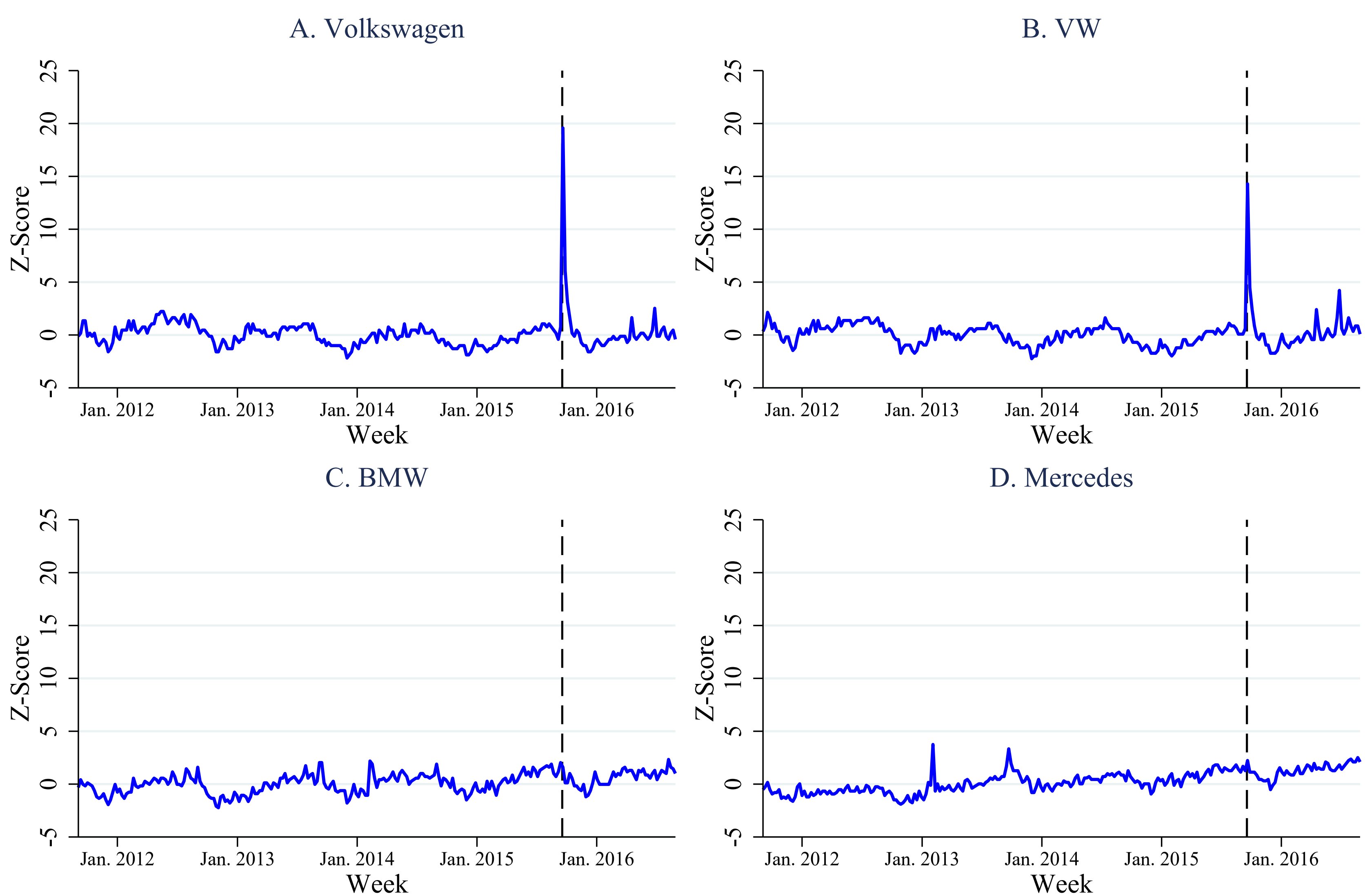
We conclude this section with a brief discussion of some scandals involving the German auto manufacturers, clarifying both that they are unrelated to the 2015 VW emissions scandal and that they erupted well after the end of our study period. In the summer of 2017, it was suggested that Mercedes-Benz had also manipulated emissions (Zeit Online, 2017), although Mercedes-Benz never admitted to wrongdoing in the United States. Note that these accusations arose almost two years after the VW scandal broke and one year after the end of our study period. Later, in the spring of 2018, BMW had its own cheating device scandal, which led to raids of its corporate offices in Germany (Ewing, 2018). However, as Ewing (2018) also reports, this affair was much smaller (11 thousand vehicles affected versus 11 million in the VW scandal). More importantly, it had no U.S. impact as none of the affected vehicles were in the United States.[29]

In 2021, the European Competition Commission found VW, BMW, and Mercedes-Benz guilty of colluding on limiting the size of diesel exhaust fluid tanks. The investigation of

[28] The underlying data on Google trends is weekly, and it is normalized by Google so that 100 corresponds to the largest number of searches per week in the search period. For weekly data, Google trends only allows users to download a few pre-defined search periods. We chose a five-year window from August 2011 to August 2016. We normalize the series and express weekly values as z-scores, deviations from the mean that are scaled by the standard deviation. A z-score of 1 indicates a 1-standard-deviation increase over the mean. Both the means and the standard deviations are constructed using the period prior to September 2015.

[29] Like Mercedes-Benz, BMW never admitted to any wrongdoing. BMW executives blamed a simple mistake that led them to install the wrong software for the implicated vehicles. The subsequent fines imposed by German authorities were $11.6 million, orders of magnitude smaller than the combined $25 billion of fines for VW in the United States and Germany (see Dobush, 2018).

Figure 6: Google Trends of Searches for German Auto Manufacturing Firms




Note: Dashed lines show the week of the Volkswagen emissions scandal, dated September 18, 2015. The underlying data on Google trends is weekly and it is scaled by Google so that 100 corresponds to the largest number of searches per week in the sample period (August 2011 to August 2016). We construct z-scores as the deviation from the mean of the series, normalized by its standard deviation. The z-scores on the vertical axis are constructed using the mean and standard deviation for each search term over the period prior to September 2015.

this case occurred after the VW emissions scandal. Moreover, “The Commission said it had no indications that the automakers coordinated with each other on the use of illegal emissions-cheating ‘defeat devices’” (BBC, 2018).

To sum up: at the time that the VW scandal broke, U.S. consumers had neither a technical nor an economic reason to believe that BMW and Mercedes-Benz were culpable of malfeasance similar to VW’s. The documented Google search behavior and Twitter sentiment expression appear to be inconsistent with the notions that BMW and Mercedes-Benz were implicated in the same way as VW, nor that U.S. consumers believed in collective malfeasance. By contrast, the evidence is consistent with our interpretation of the spillover effect as arising from collective reputation.

# 6 Conclusion

This paper documents that firms have economically important collective reputations in the context of a key industry featuring national powerhouse companies. Using the 2015 VW emissions scandal as a natural experiment, we show that these collective reputations can give rise to reputation spillovers with large economic consequences; in particular, a scandal implicating one group member can have adverse impacts on the other group members not implicated in the wrongdoing. We begin by using a difference-in-differences approach to show that the scandal reduced sales at the other German auto manufacturers relative to their non-German counterparts. The net total effect of the scandal is a combination of two countervailing forces, the substitution effect and the spillover effect. We estimate a model of vehicle demand to decompose the scandal’s effect into its individual channels. Our results suggest that the scandal had an economically important country-specific reputational spillover. We thus provide empirical support for the existence both of collective reputation for firms and of group reputation externalities and quantify their economic significance.

We believe that the existence of meaningful collective reputations among firms is likely to have practical policy implications. For instance, regulators may wish to consider collective reputations when evaluating potential mergers. Individual firms sharing collective reputations have an incentive to “free-ride” on the group’s reputation by under-investing in quality. A merger among firms sharing a group reputation may produce a procompetitive benefit by inducing the combined firm to internalize the collective reputation externality, thereby incentivizing higher investment in quality. Therefore, in scenarios where the traditional anti-competitive and pro-competitive merger effects are otherwise balanced, regulators may want to account for this novel merger effect. Although examining such policy implications falls outside the scope of this paper, we hope that our results encourage future research on this topic.

# References

Alé-Chilet, Jorge, Cuicui Chen, Jing Li, and Mathias Reynaert (2021), "Colluding against environmental regulation." Working Paper 16038, Center for Economic Policy Research.

Alexander, Diane and Hannes Schwandt (2019), "The impact of car pollution on infant and child health: Evidence from emissions cheating." WP 2019-04, Federal Reserve Bank of Chicago.

Angrist, Joshua and Alan Krueger (1999), "Empirical strategies in labor economics." In *Handbook of Labor Economcis* (Orley Ashenfelter and David Card, eds.), volume 3, chapter 23, 1277–1366, Elsevier.

Ater, Itai and Nir Yoseph (2020), "The impact of environmental fraud on the used car market: Evidence from dieselgate." Working Paper 12899, Center for Economic Policy Research.

Bachmann, Ruediger, Gabriel Ehrlich, Ying Fan, and Dimitrije Ruzic (2019), "Firms and collective reputation: A study of the Volkswagen emissions scandal." Working Paper 26117, National Bureau of Economic Research.

Bai, Jie, Ludovica Gazze, and Yukun Wang (2021), "Collective reputation in trade: Evidence from the chinese dairy industry." *Review of Economics and Statistics*, 1–45.

Barrage, Lint, Eric Chyn, and Justine Hastings (2020), "Advertising and environmental stewardship: Evidence from the BP oil spill." *American Economic Journal: Economic Policy*, 12, 33–61.

Barth, Florian, Christian Eckert, Nadine Gatzert, and Hendrik Scholz (2019), "Spillover effects from the Volkswagen emissions scandal: An analysis of stock, corporate bond, and credit default swap markets." Working Paper, Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg.

BBC (2018), "BMW, Daimler and Volkswagen face EU diesel emissions probe." *BBC*. https://www.bbc.com/news/business-45558588.

Berry, Steven (1994), "Estimating discrete-choice models of product differentiation." *RAND Journal of Economics*, 25, 242–262.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile prices in market equilibrium." *Econometrica*, 63, 841–890.

Berry, Steven, James Levinsohn, and Ariel Pakes (2004), "Differentiated products demand systems from a combination of micro and macro data: The new car market." *Journal of Political Economy*, 112, 68–105.

Breitinger, Matthias (2018), "Der Abgasskandal." *Zeit Online*. http://www.zeit.de/wirtschaft/diesel-skandal-volkswagen-abgase.

Bruckner, Johanna (2015), "Was die Skandale für das deutsche Image bedeuten." *Sueddeutsche.de*. http://www.sueddeutsche.de/kultur/pegida-dfb-und-vw-was-die-skandale-fuer-das-deutsche-image-bedeuten-1.2732828.

Cabral, Luís and Ali Hortaçsu (2010), "The dynamics of seller reputation: Evidence from eBay." *Journal of Industrial Economics*, 58, 54–78.

Castriota, Stefano and Marco Delmastra (2014), "The economics of collective reputation: Evidence from the wine industry." *American Journal of Agricultural Economics*, 97, 469–489.

Chambers, Madeline (2015), "VW scandal threatens 'Made in Germany' image." *Reuters*. https://www.reuters.com/article/usa-volkswagen-germany-image-idUSL5N11S38X20150922.

Che, Xiaogang, Hajime Katayama, and Peter Lee (2018), "Willingness to pay for brand reputation: Lessons from the Volkswagen emissions scandal." Working Paper, Durham University.

Dobush, Grace (2018), "BMW will pay $11.6 million in dieselgate fines—a tiny penalty compared to other emission scandals." *Fortune Magazine*. http://fortune.com/2018/09/03/bmw-dieselgate-fine/.

Ewing, Jack (2018), "BMW offices raided by authorities in emissions-cheating investigation." *The New York Times*. https://www.nytimes.com/2018/03/20/business/energy-environment/bmw-diesel-emissions.html.

Fan, Ying, Jiandong Ju, and Mo Xiao (2016), "Reputation premium and reputation management: Evidence from the largest e-commerce platform in China." *International Journal of Industrial Organization*, 46, 63–76.

Fiordelisi, Franco, Maria-Gaia Soana, and Paola Schwizer (2014), "Reputational losses and operational risk in banking." *European Journal of Finance*, 20, 105–124.

Fishman, Arthur, Israel Finkelstein, Avi Simhon, and Nira Yacouel (2018), “Collective brands.” *International Journal of Industrial Organization*, 59, 316–339.

Freedman, Seth, Melissa Kearney, and Mara Lederman (2012), “Product recalls, imperfect information, and spillover effects: Lessons from the consumer response to the 2007 toy recalls.” *Review of Economics and Statistics*, 94, 499–516.

Gandhi, Amit and Jean-François Houde (2019), “Measuring substitution patterns in differentiated products industries.” Working Paper 26375, National Bureau of Economic Research.

Goolsbee, Austan and Amil Petrin (2004), “The consumer gains from direct broadcast satellites and the competition with cable tv.” *Econometrica*, 72, 351–381.

Grieco, Paul, Charles Murry, Joris Pinkse, and Stephan Sagl (2021), “Efficient estimation of random coefficients demand models using product and consumer datasets.” Working paper, Pennsylvania State University.

Grieco, Paul, Charles Murry, and Ali Yurukoglu (2021), “The evolution of market power in the U.S. auto industry.” Working Paper, Pennsylvania State University.

Griffin, Paul and David Lont (2018), “Game changer? The impact of the VW emission-cheating scandal on the interrelation between large automakers’ equity and credit markets.” *Journal of Contemporary Accounting & Economics*, 14, 179–196.

Improbable Research (2016). *Improbable Research*. http://www.improbable.com/ig/winners/.

Jonsson, Stefan, Henrich Greve, and Takako Fujiwara-Greve (2009), “Undeserved loss: The spread of legitimacy loss to innocent organizations in response to reported corporate deviance.” *Administrative Science Quarterly*, 54, 195–228.

Levin, Jonathan (2009), “The dynamics of collective reputation.” *The B.E. Journal of Theoretical Economics (Contributions)*, 9, Article 27.

Li, Jing (2019), “Compatibility and investment in the U.S. electric vehicle market.” Working Paper, MIT.

Li, Lingfang (2010), “Reputation, trust, and rebates: How online auction markets can improve their feedback mechanisms.” *Journal of Economics & Management Strategy*, 19, 303–331.

Li, Lingfang, Steven Tadelis, and Xiaolan Zhou (2020), "Buying reputation as a signal of quality: Evidence from an online marketplace." *RAND Journal of Economics*, 51, 965–988.

Luca, Michael (2016), "Reviews, reputation, and revenue: The case of Yelp.com." HBS Working Paper 12-016.

Mayzlin, Dina, Yaniv Dover, and Judith Chevalier (2014), "Promotional reviews: An empirical investigation of online review manipulation." *American Economic Review*, 104, 2421–2455.

Murry, Charles and Yiyi Zhou (2020), "Consumer search and automobile dealer colocation." *Management Science*, 66, 1909–1934.

Neeman, Zvika, Aniko Öry, and Jungju Yu (2019), "The benefit of collective reputation." *The RAND Journal of Economics*, 50, 787–821.

Newburry, William (2012), "Waving the flag: The influence of country of origin on corporate reputation." In *The Oxford Handbook of Corporate Reputation* (Timothy Pollock and Michael Barnett, eds.), Oxford Handbooks Online, Oxford University Press.

Nienaber, Michael (2015), "Volkswagen could pose bigger threat to German economy than Greek crisis." *Reuters*. http://www.reuters.com/article/us-usa-volkswagen-germany-economy-idUSKCN0RN27S20150923.

Nosko, Chris and Steven Tadelis (2015), "The limits of reputation in platform markets: An empirical analysis and field experiment." Working Paper, University of California at Berkeley.

Petrin, Amil (2002), "Quantifying the benefits of new products: The case of the minivan." *Journal of Political Economy*, 110, 705–729.

Petroff, Alanna (2015), "What Volkswagen means to the German economy." *CNN Money*. http://money.cnn.com/2015/09/22/news/economy/volkswagen-germany-cars-economy/.

Pitas, Costas (2017), "Mini boss says UK production not essential to brand." *Reuters*. https://www.reuters.com/article/us-autoshow-geneva-mini-britain/mini-boss-says-uk-production-not-essential-to-brand-idUSKBN16E1F0.

Remsky, Sarah (2017), “Eine Belastung für das Label ‘Made in Germany’.” *Zeit Online*. http://www.zeit.de/wirtschaft/unternehmen/2017-07/abgasskandal-deutsche-wirtschaft-krise-volkswagen-daimler-porsche-automobilindustrie.

Springel, Katalin (2020), “Network externality and subsidy structure in two-sided markets: Evidence from electric vehicle incentives.” Working Paper, Georgetown University.

Strittmatter, Anthony and Michael Lechner (2020), “Sorting in the used-car market after the Volkswagen emission scandal.” *Journal of Environmental Economics and Management*, 101, 102305.

Thompson, Gregory, Daniel Carder, Marc Besch, Arvind Thiruvengadam, and Hemanth Kappanna (2014), “In-use emissions testing of light-duty diesel vehicles in the united states.” Report, Center for Alternative Fuels, Engines & Emissions, West Virginia University.

Tirole, Jean (1996), “A theory of collective reputations (with applications to the persistence of corruption and to firm quality).” *Review of Economic Studies*, 63, 1–22.

Train, Kenneth E and Clifford Winston (2007), “Vehicle choice behavior and the declining market share of U.S. automakers.” *International Economic Review*, 48, 1469–1496.

Werz, Michael (2016), “Smarter Lobbyieren.” *Zeit Online*. http://www.zeit.de/wirtschaft/unternehmen/2016-01/automobilindustrie-usa-volkswagen-audi-porsche-verfahren.

Winfree, Jason A and Jill J McCluskey (2005), “Collective reputation and quality.” *American Journal of Agricultural Economics*, 87, 206–213.

Xing, Jianwei, Benjamin Leard, and Shanjun Li (2021), “What does an electric vehicle replace?” *Journal of Environmental Economics and Management*, 107, 102432.

Zeit Online (2017), “Daimler soll mehr als eine Million Fahrzeuge manipuliert haben.” *Zeit Online*. http://www.zeit.de/wirtschaft/unternehmen/2017-07/abgasskandal-daimler-ag-schadstoffe?

Zycher, Benjamin (2017), “The Volkswagen emissions scandal and the urge for collective punishment.” *American Enterprise Institute Spotlight*. http://www.aei.org/spotlight/the-volkswagen-emissions-scandal/.

# A List of Automotive Makes

Table A.1: List of Makes in the Data

| | | |
|---|---|---|
| Acura | Hummer | Nissan |
| Alfa Romeo | Hyundai | Pontiac |
| Audi | Infiniti | Porsche |
| BMW | Jaguar | Ram |
| Buick | Jeep | Saab |
| Cadillac | Kia | Saturn |
| Chevrolet | Land Rover | Scion |
| Chrysler | Lexus | Smart |
| Dodge | Lincoln | Subaru |
| Fiat | Mazda | Suzuki |
| Ford | Mercedes-Benz | Tesla |
| Genesis | Mercury | Toyota |
| GMC | Mini | Volkswagen |
| Honda | Mitsubishi | Volvo |

Note: Genesis, Hummer, Mercury, Pontiac, and Saturn are not in our difference-in-differences analysis because they were either discontinued before or in January 2011, the first month of our DID analysis, or, in the case of Genesis, were launched in November 2015, after the scandal.

Table A.2: Makes in Twitter Data

| | | |
|---|---|---|
| Acura | Honda | Mini |
| Audi | Hyundai | Mitsubishi |
| BMW | Infiniti | Nissan |
| Buick | Jaguar | Porsche |
| Cadillac | Jeep | Ram |
| Chevrolet | Kia | Scion |
| Chrysler | Land Rover | Smart |
| Dodge | Lexus | Subaru |
| Fiat | Lincoln | Toyota |
| Ford | Mazda | Volkswagen |
| GMC | Mercedes-Benz | Volvo |

Note: The data come from Networked Insights.

# B Robustness: Difference-in-Differences Estimation

In this appendix, we first show that our treatment of Mini as a non-German make in the baseline specification does not drive our results. We then provide additional visual evidence to support the notion of parallel trends in sales growth rates between the non-VW German and non-German auto manufacturers prior to the scandal. Finally, we explore alternative econometric specifications for our difference-in-differences analysis.

## B.1 Classification of Mini

In our baseline country classification of auto makes, we include Mini—a company with historical roots in Britain that is now owned by BMW—as a non-German make. Our classification is supported by BMW board member Peter Schwarzenbauer, who told Reuters in a 2017 interview that the "brand being perceived as British, that's important... Most people don't know where the cars are produced" (Pitas, 2017). This focus on the country of brand association rather than the country of production or ownership drives our baseline classification choice. Nonetheless, we show here that this choice does not impact our results. Column (2) of Table B.1 excludes Mini from the analysis altogether; the resulting estimate of a 9.3 percentage point decline in non-VW German vehicle sales growth hardly differs

Table B.1: Difference-in-Differences Estimates—Robustness to Classification of Mini

| Dependent Variable | 12-month Log Sales Growth | | |
|---|---|---|---|
| Treatment of Mini | Baseline (considered non-German) | Excluded from the Sample | Considered German |
| | (1) | (2) | (3) |
| German × Post-Scandal | -0.092*** | -0.093*** | -0.099*** |
| | (0.030) | (0.030) | (0.029) |
| Time Fixed Effects | Yes | Yes | Yes |
| Make Fixed Effects | Yes | Yes | Yes |
| $R^2$ | 0.303 | 0.305 | 0.304 |
| N | 2150 | 2082 | 2150 |

Note: The unit of observation is a make-month (e.g., the log growth of all BMW sales from January 2014 to January 2015). The time period covered is January 2011 to August 2016. Standard errors are clustered at the make level and are reported in parentheses. VW Group (VW, Audi, and Porsche) is excluded from all regressions. The VW emissions scandal is dated September 18, 2015. Sales are measured in units sold. All regressions are weighted by the square root of sales volumes. The data come from WardsAuto. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

from the baseline estimate of 9.2 percentage points. Classifying Mini as German through its ownership by BMW in Column (3) leads to an estimated decline of 9.9 percentage points, which is also similar to the baseline result.

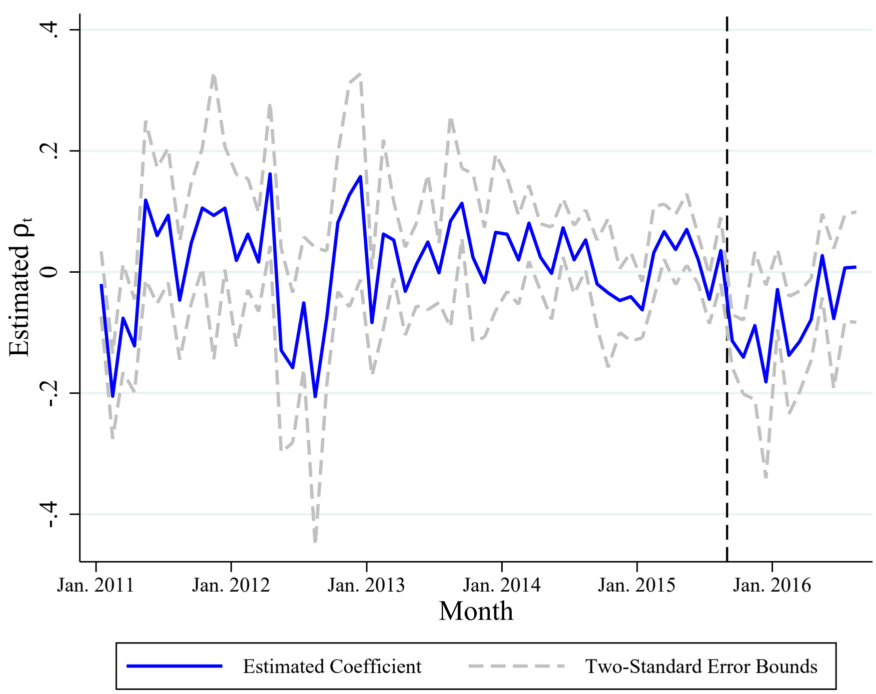
## B.2 Assessing Parallel Pre-Trends

To examine whether the trends in sales growth displayed in Figure 5 are parallel prior to the scandal, we estimate month-by-month differences in the sales growth rates of German and non-German auto manufacturers. Specifically, we estimate a regression of the form:

$$\ln \mathit{Sales}_{kt} - \ln \mathit{Sales}_{kt-12} = \gamma_t + \rho_t \mathit{non\text{-}VW\ German}_k + \varepsilon_{it}, \tag{B.2.1}$$

where $\gamma_t, t = 1, ..., T$ is a set of month dummies for January 2011 to August 2016; *non-VW German*$_i$ is a dummy variable that equals one if the make $i$ is BMW, Mercedes-Benz, or Smart; and $\rho_t$ is the difference in sales growth between German and non-German auto manufacturers, which we allow to vary over time. Figure B.2 displays the estimated coefficients $\rho_t$ and the corresponding two-standard-error confidence bands based on standard errors clustered at the make level. These estimates are centered at zero prior to the scandal, indicating that the vehicle sales of the non-VW German manufacturers grew at a

Figure B.2: Differences in Sales Growth for non-VW German vs. non-German Automakers




Note: The dashed vertical line shows the month of the VW emissions scandal, September 2015. The solid line displays the estimated coefficients $\rho_t$ from equation (B.2.1). Confidence bands are calculated from standard errors clustered at the make level. The regression excludes the VW Group. The data come from WardsAuto.

similar rate to the non-German manufacturers' sales prior to the scandal. In addition, Figure B.2 shows that, after the scandal, the sales of the non-VW German auto manufacturers grew more slowly than the sales of the non-German auto manufacturers, consistent with the difference-in-differences estimates in the main text.

## B.3 Alternative Econometric Specifications

We show in this appendix that our difference-in-differences estimates are not sensitive to alternative implementations. In Section 3.1, we weighted observations by the square root of the make's monthly sales volume. Column (2) of Table B.3 shows that our choice to weight the observations is conservative: the unweighted estimate of the difference-in-differences coefficient is 14 percentage points. In Column (3), we consider mid-point growth rates, in which the change in sales volume between period $t$ and period $t-12$ is divided by the average level of sales in the two periods. The estimated difference-in-differences coefficient under this alternative measure is 9.4 percentage points, which is similar to the baseline result in Column (1). Column (4) includes only one year of pre-scandal growth rates to match the length of the post-scandal period; limiting the pre-scandal period in this way leaves the key estimate effectively unchanged. Column (5) presents results using the level of log sales as the dependent variable. Our baseline specification uses 12-month log sales growth as the dependent variable to account parsimoniously for potential make-specific trends and make-month-specific seasonality in the level of monthly vehicle sales. In the same spirit, Column (5) also includes make-specific fixed effects for each month of the year and make-specific time trends. Our results are robust to this alternative specification.

Table B.3: Difference-in-Differences Estimates—Alternative Specifications

| Dependent Variable | 12-month Sales Growth | | | | Log Sales |
|---|---|---|---|---|---|
| Specification | Baseline | Unweighted | Mid-Point | Short Pre-Scandal Period | |
| | (1) | (2) | (3) | (4) | (5) |
| German × Post-Scandal | -0.092*** | -0.140** | -0.094*** | -0.079** | -0.096*** |
| | (0.030) | (0.058) | (0.030) | (0.034) | (0.028) |
| Time Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Make Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Make 12-Month Fixed Effects | No | No | No | No | Yes |
| Make Time Trends | No | No | No | No | Yes |
| $R^2$ | 0.303 | 0.131 | 0.322 | 0.449 | 0.994 |
| N | 2,150 | 2,150 | 2,150 | 754 | 2,556 |

Note: The unit of observation is a make-month. The time period covered is January 2011 to August 2016. Standard errors clustered at the make level in parentheses. VW Group (VW, Audi, and Porsche) is excluded from all regressions. The VW emissions scandal is dated September 18, 2015. Sales are measured in units sold. Regressions in Columns (1), (3), (4), and (5) are weighted by the square root of sales volumes. Regressions in Columns (1), (2), and (4) use 12-month log sales growth as the dependent variable; Column (3) uses mid-point sales growth; Column (5) uses log sales. Column (4) includes only one year of pre-scandal growth rates. All regressions include make and time fixed effects. The data come from WardsAuto. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

# C Additional Details on Demand Estimation & Simulations

In this appendix, we present additional details on the estimation of the model of vehicle demand from Section 4. First, we provide summary statistics for the list prices and vehicle characteristics used in the estimation. Second, we present the first-stage regression results, in which we regress the endogenous variable, vehicle prices, on the instrumental variables. Third, we present alternative decompositions of the three channels by which the VW emissions scandal affected the vehicle sales of the other German auto manufacturers. Fourth, we repeat the baseline simulations allowing for price adjustments.

## C.1 Summary Statistics of Product Prices and Characteristics

Our estimation sample consists of 1,968 product-year combinations with sales of at least 100 units, where a product is defined as a make-model-power type. Table C.1 reports the summary statistics of list prices and product characteristics.

Table C.1: Summary Statistics of Product Prices and Characteristics

| | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| Prices (2015-$1000) | 36.31 | 20.87 | 9.90 | 202.43 |
| Horsepower/Weight | 0.06 | 0.02 | 0.02 | 0.19 |
| Weight (1000 lb) | 3.89 | 0.88 | 1.81 | 7.32 |
| Miles per Gallon | 29.94 | 12.30 | 10.00 | 119.00 |
| Diesel | 0.07 | 0.26 | 0 | 1 |
| Gas | 0.79 | 0.40 | 0 | 1 |
| Other Power Type | 0.13 | 0.34 | 0 | 1 |
| Germany | 0.23 | 0.42 | 0 | 1 |
| Japan | 0.33 | 0.47 | 0 | 1 |
| Korea | 0.06 | 0.24 | 0 | 1 |
| United States | 0.30 | 0.46 | 0 | 1 |
| Other Country | 0.07 | 0.25 | 0 | 1 |
| N | 1,968 | | | |

Note: The unit of observation is a product-year. A product is defined as a make-model-power type combination (e.g., a Honda Civic with a gasoline engine). The time period covered is 2010 to 2016, with 2015 omitted as the scandal year, just as in Table 2, which reports our demand estimates. Each calendar year is considered a distinct market. The data come from WardsAuto.

## C.2 First-Stage Regression

Table C.2: First-Stage Regression for Demand Estimation

| Dependent Variable | Price ($10,000) | |
|---|---|---|
| Trim Aggregation | Min | Median |
| IV1 | -0.005*** | -0.007*** |
| | (0.000) | (0.001) |
| IV2 | 0.001*** | 0.002*** |
| | (0.000) | (0.0005) |
| IV3 | -0.0006* | -0.001* |
| | (0.000) | (0.0004) |
| IV4 | -0.0001*** | -0.001*** |
| | (0.000) | (0.0005) |
| *ln* Weight | 5.350*** | 6.547*** |
| | (0.555) | (0.630) |
| Horsepower/Weight (HP/1000lb) | 0.982*** | 1.183*** |
| | (0.031) | (0.032) |
| *ln* Miles-Per-Gallon | 0.867*** | 1.089*** |
| | (0.196) | (0.220) |
| VW Group × Post-Scandal | 0.008 | 0.243 |
| | (0.209) | (0.228) |
| Other German × Post-Scandal | 0.541*** | 0.595*** |
| | (0.180) | (0.195) |
| Diesel × Post-Scandal | 0.402 | 0.595*** |
| | (0.271) | (0.295) |
| Country Fixed Effects | yes | yes |
| Power-Type Fixed Effects | yes | yes |
| F-Statistic on Excluded Instruments | 43.095 | 73.490 |

Note: Excluded instrumental variables (IV1 – IV4) are the sum of product characteristics over closely competing products. Let $x_{jt}^{(\ell)}$ denote the value of characteristic $\ell$ for product $j$ at time $t$, then $IV_\ell = \sum_{j' \neq j} 1 \left( \left| x_{j't}^{(\ell)} - x_{jt}^{(\ell)} \right| < D_\ell \right) x_{j't}^{(\ell)}$, where for $\ell = 1, 2, 3, 4$ represents the horsepower, weight, length, miles-per-gallon of product $j$, and $D_\ell$ is the standard deviation of attribute $\ell$ in the data. Country Fixed Effects comprises indicators for German, Japanese, Korean, U.S., and All Other country origins. The data come from WardsAuto. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

## C.3 Additional Simulations

In Section 4.3, we report three simulations in which we turn on the scandal's three channels sequentially in order to quantify the effect of each. There, the sequence in which the channels are turning on is "Substitution Away from VW Group," followed by "Substitution Away from Diesel," and finally "Spillover to Other German Automakers." That sequence is, however, not unique: in fact, there are four quantifications of each channel's effect. Given the nonlinearity of the demand model, in principle these different quantifications can provide different answers regarding the magnitudes of each channel's effect. We can quantify the effect of each channel by comparing two scenarios: one in which the channel is turned on, and one in which it is turned off. For each comparison, the other two channels are either both turned on, both turned off, or one is turned on and one is turned off. Table C.3.1 summarizes the simulation designs used in these four quantifications, and Table C.3.2 reports the results. Bold table entries correspond to the baseline quantification in Table 4 in Section 4.3, and below each entry is a 95% confidence interval. As noted, the results turn out to be stable across these different quantifications.

Table C.3.1: Simulation Designs

| | S1 | S2 | S3 | Data | S4 | S5 | S6 | S7 |
|---|---|---|---|---|---|---|---|---|
| # of Active Channels | 3 | 2 | 1 | 0 | 2 | 2 | 1 | 1 |
| Substitution Effect | yes | no | no | no | yes | yes | yes | no |
| Diesel Effect | yes | yes | no | no | no | yes | no | yes |
| Spillover Effect | yes | yes | yes | no | yes | no | no | no |

Table C.3.2: Quantifications of Each Channel's Effect on the non-VW German Automakers

| | Other Two Channels | | | |
|---|---|---|---|---|
| | On/On | On/Off | On/Off | Off/Off |
| Substitution Effect | **S1 – S2** | S4 – S3 | S5 – S7 | S6 – Data |
| Vehicle Sales | **77,342** | 76,672 | 87,149 | 86,258 |
| | (26301, 130436) | (27059, 122054) | (32591, 134349) | (33817, 125245) |
| Revenue ($ billion) | **3.44** | 3.41 | 3.82 | 3.79 |
| | (1.18, 5.73) | (1.22, 5.36) | (1.44, 5.88) | (1.50, 5.48) |
| Diesel Effect | S1 – S4 | **S2 – S3** | S5 – S6 | S7 – Data |
| Vehicle Sales | 1,970 | **1,300** | 3,230 | 2,338 |
| | (–5951, 30019) | (–3810, 22561) | (–13221, 33517) | (–8545, 25607) |
| Revenue ($ billion) | 0.09 | **0.06** | 0.14 | 0.10 |
| | (–0.26, 1.31) | (–0.16, .98) | (–0.58, 1.46) | (–0.37, 1.11) |
| Spillover Effect | S1 – S5 | S2 – S7 | S4 – S6 | **S3 – Data** |
| Vehicle Sales | –255,452 | –245,645 | –254,193 | **–244,607** |
| | (–433418, –6759) | (–413989, –6653) | (–415615, –6691) | (–398503, –6590) |
| Revenue ($ billion) | –11.67 | –11.28 | –11.62 | **–11.24** |
| | (–19.83, –0.31) | (–19.03, –0.30) | (–19.03, –0.31) | (–18.34, –0.30) |

Note: Bold table entries correspond to the baseline decomposition in Table 4. Positive numbers indicate that the channel increased vehicle sales or revenue for the non-VW German automakers; negative numbers mean the opposite. The entries in the column titled "On/On" report the channel's effect in a simulation with the other two channels; e.g., the first row of that column reports the effect of the substitution away from VW channel if the scandal had also led to substitution away from diesel and a spillover to the other German automakers. The entries in the columns titled "On/Off" report the channel's effect in a simulation with one of the other two channels, as indicated by the simulation designs in Table C.3.1. Finally, the entries in the column titled "Off/Off" report the channel's effect in a simulation without the other two channels. The simulated effects are relative to 2014, the calendar year before the scandal. Revenue effects are expressed in 2015 dollars. The ranges in parentheses report 95% confidence intervals.

## C.4 Simulations with Price Adjustments

In the main simulations in Section 4, we hold prices fixed in order to quantify how much each individual channel of the VW emissions scandal shifts demand. We find that the spillover effect is significant and is larger than the scandal's net overall effect.

In this section, we describe how we allow for equilibrium price adjustment in the additional simulations in Table 5. We model the supply side of the U.S. light vehicle market, back out each product's marginal cost, and, in each simulation, compute the pricing equilibrium based on these estimated marginal costs. We assume that firms compete in a Bertrand fashion. Each firm $k$ chooses the price of each of its products, denoted by $\mathcal{J}_{kt}$, to maximize profits:

$$\max_{p_{jt}, j \in \mathcal{J}_{kt}} \sum_{j \in \mathcal{J}_{kt}} (p_{jt} - mc_{jt}) s_{jt}(\boldsymbol{p}_t), \tag{C.4.1}$$

where the vectors $\boldsymbol{p}_t$ capture the prices of all products in the market and the market share function $s_{jt}(\boldsymbol{p}_t)$ is given by the demand model in Section 4. Non-price variables such as product characteristics and demographics are absorbed in the subscript $jt$.

This profit-maximization problem implies the following first-order conditions:

$$s_{jt} + \sum_{j' \in \mathcal{J}_{kt}} (p_{j't} - mc_{j't}) \frac{\partial s_{j't}}{\partial p_{jt}} = 0 \text{ for any } j \in \mathcal{J}_{kt}, \tag{C.4.2}$$

which can be rewritten in matrix form using the following notation. Let $s_{kt} = (s_{jt}, j \in \mathcal{J}_{kt})$ and define $p_{kt}$ and $mc_{kt}$ analogously. Let $\#\mathcal{J}_{kt}$ denote the number of products sold by firm $k$ in year $t$ and $\nabla_{kt}$ be a $\#\mathcal{J}_{kt}$-by-$\#\mathcal{J}_{kt}$ matrix whose $(j, j')$ element is $\frac{\partial s_{j't}}{\partial p_{jt}}$. Then, the first-order conditions in (C.4.2) can be rewritten as:

$$s_{kt} + \nabla_{kt}(p_{kt} - mc_{kt}) = 0. \tag{C.4.3}$$

We can then solve for the marginal costs as:

$$mc_{kt} = p_{kt} - \nabla_{kt}^{-1} s_{kt}, \tag{C.4.4}$$

where market shares $s_{kt}$ and prices $p_{kt}$ are observable, while the "slopes" of the market share functions $\nabla_{kt}$ are taken from the estimated demand model.

Based on these marginal costs, we repeat each simulation in Section 4 by first computing the Nash-Bertrand equilibrium prices with some or all of the scandal's channels turned on and then calculating the corresponding sales quantities and revenues. Table 5 shows that our key findings are robust to allowing for equilibrium price changes.

# D Robustness: Demand Estimation

In this appendix, we show that the demand estimation in Section 4.2 is robust to different specifications of fixed effects. Table D.1 reports the results. Relative to the specifications reported in the main text in Table 2—copied below as Columns (1) and (4)—we now also include a combination of year and make fixed effects. Columns (2), (3), (5), and (6) all have year fixed effects. Columns (3) and (6) additionally replace country fixed effects with make fixed effects.

Given the two-step nature of the estimation described in Section 4.2, all interactions of household characteristics and vehicle characteristics are estimated first and hence are invariant to the subsequent choice of vehicle characteristics and fixed effects. Therefore, Columns (1)-(3) share the same coefficients for those interactions, as do Columns (4)-(6).

Table D.1: Robustness: Demand Estimation Results—U.S. Light Vehicle Sales

| | Trim Aggregation: Min | | | Trim Aggregation: Median | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Price (\$10,000) | -2.439*** | -2.245*** | -2.116*** | -1.932*** | -1.810*** | -1.772*** |
| | (0.104) | (0.071) | (0.055) | (0.062) | (0.047) | (0.038) |
| Price $\times$ Income (\$1m) | 3.807*** | 3.807*** | 3.807*** | 3.531*** | 3.531*** | 3.531*** |
| | (0.060) | (0.060) | (0.060) | (0.070) | (0.070) | (0.070) |
| Price $\times$ Income$^2$ | -1.382*** | -1.382*** | -1.382*** | -1.729*** | -1.729*** | -1.729*** |
| | (0.092) | (0.092) | (0.092) | (0.141) | (0.141) | (0.141) |
| Price $\times$ Age | 1.096*** | 1.096*** | 1.096*** | 0.917*** | 0.917*** | 0.917*** |
| | (0.017) | (0.017) | (0.017) | (0.017) | (0.017) | (0.017) |
| *ln* Weight | 5.319*** | 4.454*** | 3.674*** | 4.721*** | 4.028*** | 3.754*** |
| | (0.444) | (0.320) | (0.258) | (0.312) | (0.255) | (0.240) |
| *ln* Weight $\times$ 1(House- | 0.176*** | 0.176*** | 0.176*** | 0.171*** | 0.171*** | 0.171*** |
| hold Size$\geq$3) | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) |
| *ln* Weight $\times$ Income | -2.055*** | -2.055*** | -2.055*** | -2.543*** | -2.543*** | -2.543*** |
| | (0.143) | (0.143) | (0.143) | (0.144) | (0.144) | (0.144) |
| *ln* Weight $\times$ Income$^2$ | -1.451*** | -1.451*** | -1.451*** | -0.544*** | -0.544*** | -0.544*** |
| | (0.199) | (0.199) | (0.199) | (0.191) | (0.191) | (0.191) |
| *ln* Weight $\times$ Age | 1.317*** | 1.317*** | 1.317*** | 1.270*** | 1.270*** | 1.270*** |
| | (0.030) | (0.030) | (0.030) | (0.043) | (0.043) | (0.043) |
| Horsepower/Weight | 0.035*** | 0.021*** | 0.009*** | 0.025*** | 0.015*** | 0.009*** |
| (HP/1000lb) | (0.006) | (0.004) | (0.003) | (0.004) | (0.003) | (0.003) |
| *ln* Miles-Per-Gallon | 1.521*** | 1.227*** | 1.452*** | 1.610*** | 1.278*** | 1.506*** |
| | (0.164) | (0.150) | (0.167) | (0.165) | (0.159) | (0.187) |
| VW Group $\times$ | -0.486*** | -0.604*** | -0.337** | -0.391** | -0.577*** | -0.380** |
| Post-Scandal | (0.159) | (0.160) | (0.151) | (0.159) | (0.176) | (0.174) |
| Other German $\times$ | -0.327* | -0.525*** | -0.744*** | -0.413** | -0.613*** | -0.770*** |
| Post-Scandal | (0.175) | (0.166) | (0.172) | (0.165) | (0.167) | (0.173) |
| Diesel $\times$ | 0.071 | -0.105 | -0.265 | 0.145 | -0.039 | -0.174 |
| Post-Scandal | (0.241) | (0.226) | (0.246) | (0.239) | (0.234) | (0.242) |
| Nest parameter | 0.605*** | 0.605*** | 0.605*** | 0.604*** | 0.604*** | 0.604*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Country Fixed Effects | yes | yes | no | yes | yes | no |
| Power-Type Fixed Effects | yes | yes | yes | yes | yes | yes |
| Year Fixed Effects | no | yes | yes | no | yes | yes |
| Make Fixed Effects | no | no | yes | no | no | yes |

Note: Columns (1)-(3) use the minimum characteristic across trims as the product characteristic, and Columns (4)-(6) use the median. A product is defined as a make-model-power type combination (e.g., a Honda Civic with a gasoline engine). The time period covered is 2010 to 2016, with 2015 omitted as the scandal year. Each calendar year is considered a distinct market. Country Fixed Effects comprises indicators for German, Japanese, Korean, U.S., and All Other country origins. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.