Did Chrysler Benefit from Government Assistance?
Making Causal Inferences in Small Samples using Synthetic Control Methodology

Adam R. Fremeth[‡], Guy L. F. Holburn[*], Brian Kelleher Richter[†]

This version: 26 September 2012

Abstract
We assess the impact of U.S. government intervention on the performance of Chrysler which, following the financial crisis in 2008, accepted government support in return for on-going Treasury oversight. To conduct our analysis, we introduce a recently developed statistical methodology—synthetic control—to management research. The method overcomes challenges to causal inference in contexts like ours that are constrained by small samples or few occurrences of the phenomenon of interest. Our synthetic control analysis constructs a replica of the focal firm based on a weighted combination of other firms with similar attributes. We quantify the magnitude and direction of the treatment effect by comparing the actual performance of Chrysler to its counterfactual replica without treatment. We estimate that Chrysler, or its surviving components, would have sold approximately 20% more vehicles in the U.S. through summer 2011 had the company instead relied on private financing and market forces.
Keywords: Synthetic Control, Counterfactual Methodology, Government Assistance, TARP, Automobile Industry, Case Study Methods

---

[‡] Adam R. Fremeth. afremeth@ivey.uwo.ca; Assistant Professor of Business, Economics, and Public Policy; Ivey School of Business, University of Western Ontario, London, ON N6A 3T9 CANADA
[*] Guy L. F. Holburn. gholburn@ivey.uwo.ca; Associate Professor of Business, Economics, and Public Policy and Suncor Chair in Energy Policy; Ivey School of Business, University of Western Ontario, London, ON N6A 3T9 CANADA
[†] Brian Kelleher Richter. brian.richter@mccombs.utexas.edu. Phone (USA): 310-709-5745. Web: http://briankrichter.com/. After January 1, 2013: Assistant Professor of Business, Government, and Society; McCombs School of Business; University of Texas at Austin; Austin, TX 78705 USA. Through December 31, 2012: Assistant Professor of Business, Economics, and Public Policy; Ivey School of Business, University of Western Ontario, London, ON N6A 3T9 CANADA

**Fremeth acknowledges support from the Social Science and Humanities Research Council of Canada (SSHRC) Grant R4636A02. Any errors in this draft are our own.

1 INTRODUCTION

Direct government involvement in the private sector has become more prevalent following the financial crisis of 2008, renewing old questions about the impact of government ownership and control on firms' performance. We examine the performance implications of government involvement in firms within the context of U.S. government intervention in the auto industry starting in late 2008. Specifically, we ask what is the impact on subsequent performance at Chrysler, or its surviving components, of accepting government financial assistance in return for ongoing Treasury oversight and restrictions on managerial actions? Implicit government guarantees of the firm as a going-concern and access to capital at below market rates may have helped Chrysler sell more vehicles. Alternatively, government involvement could have led to fewer sales of Chrysler vehicles if, as a result, consumers spurned the "bailed out" firm (Davis & Thompson 1994), executive talent became difficult to retain (Jensen & Murphy 1990; Pfeffer 1994), conflicting public-private objectives led to suboptimal decision making (Priem 1990; Eisenhardt & Bourgeios 1988), or competitors bolstered their efforts (Seamans 2012). In our analysis—using synthetic control methodology which we introduce to management research—we estimate that Chrysler sold 20% fewer vehicles than it would have without government involvement, suggesting that any managerial benefits of government guarantees or cheap capital were outweighed by other potential costs.

Media personalities and politicians have debated the broader topic widely, concocting competing stories for why government intervention in the auto industry either succeeded or failed (Romney 2012; Stewart 2012).[1] A *New York Times* columnist suggested that despite the merits of the debate it would be impossible to construct the idealized counterfactual necessary to conduct a rigorous analysis of the auto assistance program, writing that "unlike a science experiment, in which variables can be changed and the experiment repeated, we can't turn back the clock, let the auto companies go bankrupt and compare the results with what we have today" (Stewart 2012). This perspective addresses the limits of non-experimental research; researchers left with observational data must make tradeoffs between accuracy, generality, and simplicity (Thorngate 1976; Weick 1979). Given the limitations of the context, predominant empirical methodologies for observational data cannot be used to answer our research question, which requires us to quantify the net effect of government assistance. Sufficiently precise quantitative estimates generated using either differences-in-differences or matched sample methods are precluded due to the small number of firms that are

[1] Republican presidential candidate, Mitt Romney (2012) famously quipped: "The president tells us that without his intervention things in Detroit would be worse. I believe that without his intervention things there would be better."

auto manufacturers and the even smaller number of firms that took government assistance. Event studies using stock market data are precluded because the focal firm, Chrysler, was privately-owned. Hence, our specific question requires a new methodology.

In this paper we apply synthetic control methodology (Abadie & Gardeazabal 2003), which has been used in economics, law, and political science research but hitherto not in management, to develop insights into the impact of government intervention on Chrysler's sales after 2008. The synthetic control technique is particularly valuable in contexts where limited sample sizes or few occurrences of the phenomenon of interest preclude researchers from making strong inferences using traditional regression techniques. The synthetic control method generates a counterfactual statistical unit, i.e. a synthetic clone of the focal unit that behaves as if it was not subject to an intervention or treatment phenomenon that has been applied to the focal unit.[2] This counterfactual unit comes complete with its own data on performance and descriptive attributes.[3] The synthetic is constructed based on a weighted average of other observation units in the population that (i) were not subject to the treatment phenomenon that is the focus of analysis, and (ii) have observable performance and descriptive data. The synthetic unit will closely mimic the focal unit in both its descriptive characteristics and pre-treatment performance. Hence, the synthetic will provide a better counterfactual than any single quantitative observation or qualitative case. Researchers can compare the relative performance of the focal observation unit and the counterfactual synthetic unit to estimate the magnitude and direction of divergence attributable to the phenomenon of interest.

The synthetic control method enables researchers to overcome shortcomings associated with existing qualitative or case-based methods typically applied in contexts like ours that are constrained by few observations. Given the limited applicability of large sample statistical methodologies to answer our question, management scholars might be left with few non-qualitative tools. A simple approach would be to conduct a single case study examining the direction of Chrysler's performance subsequent to accepting government assistance; however, that analysis would be flawed as it lacks any counterfactual. Hence, comparative case analysis would be a more appropriate tool. In that approach, a researcher would want to identify the most similar automobile company that did not

[2] Another statistical technique that attempts to construct a representative counterfactual from existing data is the 'chop shop' method used to estimate the diversification discount (Lang & Stulz 1994; Villalonga 2004; Laeven & Levine 2007).

[3] Both synthetic control and simulation methods generate data; however, they differ in that in synthetic control researchers need not make assumptions about the underlying data-generating process through formal mathematical models of agent behavior. Rather in synthetic control, the core assumption about the data generating process is that a treated unit shares the same data-generating process as control units with like attributes.

accept government assistance to use as the counterfactual evidence.[4] This comparative case could be used to draw conclusions about the direction, and size, of the effect of Chrysler's acceptance of government assistance. If we were to select the best counterfactual for Chrysler among existing firms, we might choose Ford as it was the only one of the Big Three U.S. automakers that did not take government assistance. A simple graphical comparison of Chrysler's and Ford's unit sales in Figure 1 suggests that Chrysler substantially underperformed Ford once it began receiving government assistance.

<Insert Figure 1 Here>

Causal claims drawn from Figure 1 would be impossible to support—even if they were limited to claims about the direction of the effect of government assistance on Chrysler and not its magnitude. Ford would be easy to critique as the only counterfactual, or control unit, for a variety of reasons: despite similarities with Chrysler, Ford sells different vehicles with different attributes at different price points, uses slightly different technology in its manufacturing process, has different technical expertise, and has a different marketing and branding program than does Chrysler. We will quantify these differences between Ford and Chrysler and show what will make an even better comparison, using synthetic control, later in this paper. Any one of the attributes that makes Ford different than Chrysler could have affected the companies' relative sales in the period after December 2008, making it hard to attribute a performance differential to Chrysler's acceptance of government assistance alone. That is unless we could compare Chrysler's performance with a firm's that was a better match than Ford.

If there were a company that matched Chrysler better along more dimensions than Ford, it would provide a more ideal comparison for evaluating the impact of government assistance on Chrysler's sales. Since other automakers may be more similar to Chrysler on some dimensions than Ford, our analysis might be more accurate were we to apply the logic of multiple case analyses (Eisenhardt 1989) to construct an average of all valid comparator companies' performance in the absence of government assistance.

Better yet, there would be a way to take a weighted average of several automakers similar to Chrysler in which we gave heavier weights to companies that are most similar along the most important dimensions for predicting Chrysler's performance. Synthetic control allows researchers to determine systematically what that weighted average should be and how important various attributes

[4] While choosing the next most similar company is one comparative case selection logic that is particularly useful when researchers are trying to make inferences about causal direction and magnitudes, another case selection logic suggests choosing the most extremely different case when trying to make inferences about underlying mechanisms.

of a company are in predicting performance. Hence, the methodology allows us to construct a counterfactual Chrysler that does not accept government assistance based on a sample of control companies that did not do so. We can then compare Chrysler's actual performance to its synthetic counterfactual's. In our application of synthetic control, we find that Chrysler underperforms its synthetic counterfactual by 20% following its acceptance of government assistance. Moreover, the underperformance is not as severe as we would have estimated it was had we used Ford as the sole counterfactual.

In the next section we discuss in more detail the operation of the synthetic control method, its usage to date in cognate disciplines' research, its advantages, and its limitations. In section 3, we demonstrate the method's value to management researchers in our application where we quantify the performance impact of a firm's decision to accept conditional government assistance. In the final sections we explore reasons that may account for the estimated shortfall in Chrysler's sales; we also discuss the limitations of our analysis while providing suggestions for future research.

2 THE SYNTHETIC CONTROL METHOD

Synthetic control methods are beginning to gain popularity in academic fields including economics, public policy, political science, and law—but have yet to take root in management.[5] It has been used to measure the effects of: (i) domestic ETA terrorism on regional growth within the Basque region of Spain (Abadie & Gardeazabal 2003); (ii) a new tobacco tax policy in California on cigarette sales (Abadie et al. 2010); (iii) reunification of Germany on the wealth of people in the former West Germany (Abadie et al. 2011); and (iv) gun control laws on crime in several U.S. states (Donohue & Aneja 2012). Synthetic control quantifies the magnitude and direction of each of these effects studied by researchers in cognate disciplines and provides visual evidence by graphing divergence between the outcomes realized by focal units and their synthetic counterfactuals after key events.

Despite the breadth of topics studied in other disciplines, all of the questions researched using synthetic control are: (i) focused around a single event or application of a treatment phenomenon; (ii) have theoretical ambiguity that can only be resolved by identifying the direction and magnitude of the treatment effect with relative precision; and (iii) cannot be answered with case analyses due to the lack of obvious stand-alone counterfactuals among the few control units. Given that management researchers often share these initial conditions, opportunities abound to apply the method in this

[5] Marx et. al (2009) is the only work we are aware of in management even referencing synthetic control; however, they use a difference-in-difference methodology for their core analysis and only mention synthetic control in passing.

discipline. Diving deeper into how the technique works, its advantages, and its limitations should make its utility for management researchers even more explicit.

2.1 Overview of the Technique

The synthetic control technique creates a counterfactual observation unit whose performance can be compared to a focal observation unit that has undergone any sort of treatment. In a managerial context, these units are industries, firms, corporate divisions, or employees.

The counterfactual observation unit, or synthetic control unit, is constructed as a weighted average of untreated comparison or control units. The technique maximizes the ability of the synthetic unit to generate outcome data as if it were the focal unit had it not been treated. It does so by using data from the pre-treatment period (i.e. the pre-event or pre-intervention period) to minimize the difference between (i) observable values of attributes of the focal unit that are determined to be good predictors of the selected outcome variable and (ii) values of the same attributes in the synthetic unit. In essence, the researcher is selecting positive or null weights on all potential control units in the pre-treatment period to create a synthetic unit that replicates, as best it can, the outcome variable in the treated unit during the pre-treatment period.

The weights on the control units, determined using pre-treatment data, can be applied to generate post-treatment outcomes for the synthetic unit. Those post-treatment outcomes can then be interpreted as if they were the counterfactual outcome values, assuming an acceptable fit can be created such that the synthetic and the focal unit track one another in the pre-treatment period. Divergence in outcome values between the synthetic and focal unit may happen in the post-treatment period if the intervention has a significant effect.

To implement the method, we first must build the pool of all potential comparison or control units—i.e. all similar statistical units that did not receive the treatment and for which data can be collected on *k* attributes which are potential predictor variables (X) for the outcome variable (Y) in question. The synthetic control technique subjects the comparison units' predictor variables' attribute data in the pre-treatment period to a dual optimization process that minimizes:

$$\sum_{m=1}^{k} v_m \left(X_{1m} - X_{0m}W\right)^2$$

by selecting the optimal values of W and v_m—where X_{1m} is the value of the *m*-th attribute of the focal unit; X_{0m} is a 1 x *j* vector containing the values of the *m*-th predictor attribute of each of the *j* potential comparison or control units; W is a vector of weights on control units; and v_m is a vector of weights on attributes of the control units such that they maximize the ability to predict the outcome variable of interest (Abadie et al. 2011). In essence, this optimization process minimizes prediction error between the actual and the synthetic in the pre-treatment period.

Y_1 is the observed outcome data for the focal, treated, unit. $Y_o W$ is the synthetically generated outcome data for the synthetic unit in both the pre- and post-treatment periods; more simply, $Y_o W$ is the weighted average of outcome variables for the included control units.

If there are no important omitted predictor variables then a reliable synthetic match will have been created such that $Y_1 - Y_o W$—or the distance between the actual unit's outcome variable and the synthetic unit's outcome variable—will be small in the pre-intervention period (Abadie et al. 2010, 2011). This is particularly likely when the pre-intervention period is sufficiently long.

Whether the value of $Y_1 - Y_o W$ remains the same size, becomes increasingly positive, or increasingly negative in the post-treatment period allows us to make inferences about the direction and the magnitude of the treatment effect. These inferences can be further validated through robustness checks, including conducting placebo interventions both at different points in time and among control units that we know were not treated in practice but that have good synthetic matches in the pre-intervention periods (Abadie et al. 2010, 2011).

Abadie et al. (2010, 2011) provide a wealth of additional technical details and proofs supporting the underlying synthetic control methodology. Among other things, Abadie et al. (2011) prove that the underlying mathematics of the synthetic control methodology collapse down to those in a traditional regression with an additional restriction that the linear combination of weights in synthetic control must sum to one whereas regression coefficients need not be restricted to doing so. Software is available in R, Matlab, and STATA to implement synthetic control methods.[6]

2.2 Advantages

The synthetic control method has a number of advantages relative to existing techniques. In some cases, synthetic control empowers researchers to answer questions that existing methods cannot feasibly address. In other situations, where existing methods work, the synthetic control method can provide improvements to estimates by overcoming challenges that could bias results.

The primary advantage of synthetic control vis-à-vis traditional regression techniques is that the method is feasible when only one observation unit within a (potentially limited) population receives the treatment. Most important managerial questions preclude experimental designs, limiting researchers to observational data that frequently contains few good direct counterfactuals, but more often none whatsoever (March et al. 2001; Runde & de Ronde's 2010). In situations such as these,

[6] Programs are available for download on Jens Hainmuller's website at: http://www.mit.edu/~jhainm/synthpage.html.

regression methods cannot be applied. Regression methods require variation in key variables across multiple observation units making estimation using those techniques infeasible in application to these types of questions. Nevertheless, in each of the situations above only one unit varies significantly from the others. Synthetic control provides a means to address the problem of only a single observation unit being treated.

Another advantage is that synthetic control can feasibly be applied to a variety of outcome variables. This stands in contrast to event study approaches, which are the only regression-based technique that can, in some circumstances, be applied to the analysis of a single treated unit. In event studies, the only possible outcome variable is abnormal financial market returns; in synthetic control the possibilities for outcome variables are only limited by the availability of data and the researchers' imagination. This is particularly important in management research since scholars are interested in broader measures of performance than abnormal financial returns. Moreover, for some firm-level questions we may need indicators applicable to private firms as well as public ones. Furthermore, we may be interested in managerial outcomes at other levels of analysis. A final advantage relative to event studies is that, using synthetic control, we observe the entire evolution of outcomes after a treatment rather than being constrained by an event-window—enabling us to assess simultaneously short-run and longer-run implications for a particular observation unit.

In addition to providing feasible ways to answer new questions, synthetic control can be used to overcome biases in existing methods, including: researchers' cognitive biases, omitted variable bias, and endogenity. Case selection may be biased by researchers' beliefs about what makes for good comparator groups (Eisenhart 1989, Eisenhart and Graebner 2007), even in multiple comparative case studies; this is not a problem in synthetic control because the matching of comparator units is left to an objective mathematical optimization process.[7] Even fixed-effects approaches in regression techniques cannot solve the problem of time-varying unit-level omitted variables bias; however, Abadie et al. (2010) demonstrate that in synthetic control—if a synthetic unit's outcome data tracks the focal unit's well over a sufficiently long pre-treatment period—the omission of unobserved variables need not be a concern.[8] Endogeneity biases are probably the biggest challenge facing researchers in any non-experimental design claiming causality; fortunately,

[7] Cognitive biases may tempt qualitative researchers into selecting comparator units that are competitors, or as a whole appear similar, to the focal unit, while the synthetic control process disciplines us into choosing comparator units with attributes that predict an outcome measure instead.

[8] The intuition is straightforward: only units that are similar in terms of observed and unobserved determinants should produce similar outcomes and trajectories over extended durations. Hence, once a synthetic is established that closely mimics the actual in outcome behavior over an extended pre-treatment period, any discrepancy that arises in the post-treatment period may reasonably be attributed to the treatment itself.

synthetic control methods alleviate these concerns both through a quasi-experimental approach (i.e. examining a treatment applied to some units but not others) and through accompanying tests that falsify non-random assignment concerns (Abadie et al. 2011).[9] Hence, synthetic control can also be used as a robustness check on other methods' findings when researchers cannot alleviate concerns about these common biases.

2.3 Limitations

Synthetic control is not a panacea despite its many advantages—and should only be used when the question and the data are well suited to it. Similarly, external validity is an issue limiting researchers' ability to extrapolate beyond the results—that is, we cannot interpret results beyond making simple comparisons between the focal unit and its synthetic.

While providing a bridge between large sample statistics and case study methods, synthetic control differs from both. Since the goal is to quantify causal effects, capturing both the direction and magnitude of interventions, synthetic control may appear similar to large sample statistical approaches. In one sense, this comparison is fair because synthetic control shares the same limitations as large sample statistical techniques for theory building. Synthetic control itself does not provide insights into the mechanisms causing any change in the outcome variable (or lack thereof) following the treatment—which is a task to which other techniques like fuzzy set qualitative case analysis might be better suited (Fiss 2007). Hence, qualitative case studies, which can describe in rich detail a range of mechanisms, are a natural complement for synthetic control analysis.

In a another sense, synthetic control is more similar to comparative case methods than large sample statistics, because an advantage of the technique is that it can be applied to narrow contexts in which a single statistical unit is treated. Nevertheless, it cannot always provide answers. In particular, the method will fail to create suitable synthetic matches for units that are outliers or have extreme (large or small) values on the outcome measure of interest.[10] Hence, qualitative case studies may still be better suited when researchers are interested in extremes.

[9] Placebo tests can show whether or not random assignment of the treatment to units other than the focal unit or to the focal unit at different points in time lead to similarly large effects that would undermine our ability to make causal inferences. (Heckman and Hotz 1989)

[10] This is because: (i) the weighting process assigns weights—that sum to the value of one—to comparator units; and (ii) there typically is a reason why the extreme units, in fact, may have a different data generating process than the other units. By construction, it is not possible to mimic the largest or smallest unit within a population, limiting the technique to more moderate or less extreme units.

3 APPLICATION: CHYSLER'S ACCEPTANCE OF GOVERNMENT ASSISTANCE

Synthetic control methodology allows researchers to analyze phenomena that occur in a limited population and/or that apply to only a small number of firms. This enables researchers to address scenarios for which existing techniques are inappropriate. One such scenario grew out of Chrysler's decision to accept conditional government assistance in the midst of the financial crisis that began in 2008. We assess the impact of that decision on Chrysler's subsequent sales. The estimates we generate using synthetic control contribute to a broader academic debate within the management literature on the role of government in private industry (Mahoney et al. 2009; Kivleniece & Quelin 2012)—in which predictions about the impact of government intervention on firm performance range the spectrum from positive to negative.

We conduct a synthetic control analysis of this scenario in part to demonstrate the applicability of the method to a management question. We implement our analysis in STATA using the Synth Package. Our data and a highly commented .do file with step-by-step commands to replicate our analysis are available for download.[11]

3.1 Background

Beginning with the New York Federal Reserve Bank's emergency loans to Bear Stearns in March 2008, capital markets in the U.S. began to tighten, leaving many firms across sectors in unexpectedly vulnerable positions. Auto manufacturers globally were hurt by this tightening as consumer vehicle purchases are particularly sensitive to the availability of cheap personal credit; moreover, the manufacturers themselves relied heavily on revolving credit to keep operations running. By early September, all three Detroit-based auto manufacturers had requested the United States Congress provide them with temporary financial relief. Aiming to alleviate burgeoning damage to the economy, Congress overcame partisan tensions to enact the Troubled Asset Relief Program (TARP) in October 2008, with the express purpose of injecting funds into the flailing financial sector. Despite the passage of TARP, the Senate and the House could not overcome differences on the proposed terms of an auto industry relief package.

Seeing no imminent legislative agreement, President Bush sidestepped Congress on December 19, 2008; he used his executive authority to order the Treasury Department to extend an initial $13.4B of TARP funds to Chrysler and General Motors, explicitly spelling out some conditions for the assistance, while leaving other conditions to the discretion of Treasury officials who would provide oversight and be responsible for future negotiations and future injections of

[11] The data and code are available for download at: http://www.briankrichter.com/synth_control_in_mgmt_chrysler.zip.

capital. President Bush's decision to extend the use of TARP funds beyond the financial sector was controversial and some claimed illegal (Sullum et al. 2009). Nevertheless, he argued that extending loans to auto manufacturers was a suitable use of TARP because it would be irresponsible to allow iconic American firms to go bankrupt and have their constituent parts sold off if that outcome would lead to the loss of American jobs (Office of the Press Secretary 2008).

Despite having initially requested relief from Congress, Ford deemed unacceptable the structure of the ultimate agreement that Chrysler and General Motors made with the Executive Branch. Ford cited overly restrictive conditions and the potential for additional restrictions as reasons for their concern. Ford also worried about consumer perceptions of taking a "bailout" (Dolan 2009).

One initial condition of the assistance program gave the Treasury "the power to block any large transactions" (Office of the Press Secretary 2008). Later negotiations between the recipient firms and the Treasury would, among other things, lead to specific limits on executive compensation and force both companies through a managed bankruptcy process to facilitate restructuring in a way such that both the U.S. government and a United Auto Worker (UAW) union trust would take equity stakes in the firm.

How General Motors and Chrysler would perform upon accepting the conditional government assistance and how the incoming Obama administration would manage the Treasury's relationship with these firms remained unknown. Would the firms have done better, worse, or about the same without government assistance? As discussed above, synthetic control offers a means to estimate what the counterfactual performance of Chrysler, or its surviving components, would have been had the firm not accepted the assistance package President Bush offered and President Obama administered.

3.2 Data

The population of firms that form the basis for constructing a synthetic Chrysler consists of all major commercial automotive firms selling into the United States from January 2001 to December 2011; there were 19 domestic and foreign manufacturers active during this period. The primary data source is *Ward's Auto*, which is an industry-specific commercial data provider that collects monthly data that is used by automakers, dealers, parts suppliers, and the financial community. It provides comprehensive coverage of all automotive firms over the panel, including items such as sales, technical specifications, production capabilities, retail pricing, brands, and fleet composition. We supplement the automotive data with corporate level information from *Compustat*.[12] The comprehensive nature of the data enables us to construct a synthetic counterfactual firm that closely matches the actual firm on most dimensions—ultimately giving us confidence in the quality of the match.

3.2.1 Outcome Variable

We focus on U.S. monthly light vehicle sales volume by firm (i.e. the total number of new cars, light trucks, and vans sold). Sales are a key top-line performance measure for the automotive industry, commonly quoted in the media and closely scrutinized by policymakers and financial analysts alike. The monthly frequency of its release makes it a good candidate for study as it provides immediate insight into the impact of the assistance program on Chrysler. Nevertheless, monthly sales data are quite noisy (as seen in Figure 1), given periodic corporate promotions and seasonal trends in consumer behavior, so we construct a centered 12-month moving average of sales as the outcome measure we use in our analysis. A final reason we use sales as a performance measure (and one reason we use synthetic control methodology), is that Chrysler was a private company at the time it accepted government assistance, precluding the use of a financial metrics and traditional event study methods.

3.2.2 Predictor Variables

We build from *Wards* several series of explanatory variables that predict light vehicle sales for any auto company. These include attributes of the vehicles sold, along with attributes of the firms such as measures of production capabilities, scope, financial condition, and past performance. The synthetic control method does not place conditions on the number of predictors required; given its dual minimization process it will assign low weights to predictor variables in construction of a synthetic if they have little explanatory power.

The first set of independent variables captures, on a monthly basis, vehicle-specific factors that could drive consumer demand, including: average price, average fuel economy, maximum fuel economy, average size of engine, and the average weight of the vehicles sold. The second set of variables capture strategic and operational factors that distinguish firms, including: the number of active production platforms; the number of active brands; the number of active series within those brands; the number of market segments in which they compete (i.e. luxury, small car, crossover, etc.); the fraction of sales that are from SUVs, light trucks, and vans; and, the fraction of sales of

[12] Since Chrysler was a privately held company over a period of the sample we consulted separate SEC filings to supplement the Compustat information for that firm on the leverage and employees data.

imported vehicles. We are able to calculate all of the above variables from *Wards* on a monthly basis. Further firm features we incorporate into our set of predictor variables are available on an annual basis only in *Compustat*; these are the total number of employees and leverage ratio (total debt/equity). Finally, we construct as additional predictors of performance measures of past performance; these include a twelve-month moving average of sales volume and the level of that moving average indexed to 100 at the time of the government loans. Appendix A provides data sources and details of the data's construction.

3.3 Analysis

Other than selecting outcome variables and predictor variables, the remaining decisions researchers must make in conducting synthetic control are choosing (i) the intervention date and (ii) the length of the pre-intervention window over which to minimize prediction error.

In our context, the intervention date is easy to select. President Bush authorized the disbursement of government funds to Chrysler in late December 2008 and President Obama began regularly monitoring their usage in early February 2009, so we use the mid-point of these two dates, January 2009, as the intervention or treatment date.[13]

Choosing an appropriate pre-treatment window requires selecting a sufficiently long window over which to minimize prediction errors. We found that the synthetic Chrysler's behavior closely matched that of the actual Chrysler when using a 48 month window.[14]

3.3.1 Core Results

Table 1 provides summary statistics of the attributes of the synthetic Chrysler we construct along with actual data on Chrysler and Ford, all in the pre-intervention period. It shows that the synthetic Chrysler compares well to the actual Chrysler in the period prior to the firm's acceptance of government assistance. It further illustrates that the synthetic Chrysler is a closer match to the actual Chrysler in this period than Ford as a comparative case. For ten of fifteen of the observable attributes, i.e. in the majority of cases, the synthetic Chrysler is a closer match to the actual Chrysler than Ford.

For those attributes in Table 1 where the distance between the attributes' value in the synthetic Chrysler and the actual Chrysler is relatively high, the predictive power of those attributes tends to be lower—as calculated in the attribute weight matrix, v_m. Note that in synthetic control analyses we do not need to impose assumptions on which predictor variables might matter most, as

[13] We also tried other months around January 2009 as a robustness check and find little difference in the results.

[14] We experimented with other windows over which to minimize prediction error and settled on 48 months as appropriate because longer windows did little to improve the fit and shorter windows often led to a worse fit. Our results are nevertheless robust to different window lengths.

sometimes must be done in comparative case analysis, because synthetic control follows an objective, data-driven process.

<Insert Table 1 Here>

Table 2 shows which other auto manufacturing firms are potential control units that could comprise the synthetic Chrysler. It also shows what weights (W) these control companies receive in construction of the synthetic Chrysler. Recall that these weights are determined by how well they replicate Chrysler's actual performance in the pre-government assistance or pre-intervention period to create a match. When these weights are applied to post-intervention data, they generate counterfactual performance as embodied in the synthetic Chrysler that did not accept government assistance.

Of the eighteen auto manufacturers other than Chrysler that are potential matches in a synthetic, two must be excluded as matches in the synthetic Chrysler for technical reasons. General Motors also accepted government assistance when Chrysler did, meaning it was subject to the same treatment at the same time, implying that its performance after the intervention at Chrysler would not be a valid counterfactual. Jaguar Land Rover does not have leverage or employee data, attributes on which we created matches; hence, it must be excluded as well.[15]

Among the remaining sixteen firms which could comprise the synthetic, four receive positive weightings and the remaining twelve receive zero weight. It is typical in applications of synthetic control that a number of potential control units, or other auto manufacturers in this case, receive zero weights because they do not make good individual matches and none of their attributes are sufficiently similar to the focal unit's.

Of the four firms that receive positive weightings—Ford, Toyota, Isuzu, and Suzuki—intuition helps reconcile the weights the companies receive. Most auto industry observers would select Ford as the closest match to Chrysler had General Motors been ruled out as an option. Moreover, Ford had initially requested help from the government but later distanced itself from that proposal over concerns regarding the imposed structure and oversight of the loan conditions. These facts make it unsurprising that Ford receives the largest weight in the synthetic at 0.655.

If industry observers had to go beyond the Big Three, many would select Toyota as the next closest match to Chrysler given its diverse product offerings, its multiple brands and production platforms, its vast network of production facilities in the United States, and its heavier weighting

[15] Due to its unique organizational structure Jaguar-Land Rover did not report separate balance sheet or operational statements. This made it infeasible to include the leverage ratio or the number of employees for this firm which prevented us from including it as a potential control unit.

towards SUVs, light trucks, and vans than most other large Asia- and Europe- based manufacturers. This again makes it unsurprising that Toyota receives a relatively high weight of 0.152 in the synthetic Chrysler.

Isuzu and Suzuki may at first seem like odd companies to include in the synthetic Chrysler, but understanding why helps illustrate an advantage of the programmatic way in which the synthetic control method selects control firms. Isuzu's and Suzuki's inclusion can be reconciled when recognizing that these smaller Asia-based manufacturers were extremely heavily weighted towards SUVs, light trucks, and vans—which were an important part of Chrysler's pre-government assistance product portfolio and which were highly sensitive to demand fluctuations.

<Insert Table 2 Here>

Using the weights on the companies in Table 2 we generate the synthetic firm's performance and compare it to Chrysler's actual performance in Figure 2. The figure illustrates how the synthetic Chrysler would have performed relative to the actual firm—from the beginning of the pre-treatment period in January 2005 through June 2011, giving us 30 months to observe the effects of government assistance on Chrysler's performance.[16]

Prior to Chrysler's acceptance of government funds in January 2009, the synthetic Chrysler and the company's actual performance track each other quite well with periods where either the synthetic or the actual company outperform each other marginally; however, the two series never stray very far from each other. There is also a downward trend in both Chrysler's and the synthetic's performance which was indicative of the entire auto industry in the several years prior to the intervention, as the U.S. economy began to slow starting in 2005 and as fuel prices began a rapid ascent in 2007. Given that the series track each other well despite the turmoil in the industry, the synthetic achieves a good match for Chrysler's actual performance.

What is most striking in Figure 2, however, is that in the post-government assistance period Chrysler significantly underperforms its synthetic counterfactual, representing performance of the firm had it forgone the government's offer of conditional assistance. This suggests that Chrysler would have been able to sell more vehicles had it not accepted the government's offer. Had the two series instead continued to track each other, in the post-intervention period, then we would not have been able to make this inference. Similarly, had the synthetic underperformed Chrysler in practice then we could conclude that government assistance increased subsequent sales.

[16] While our dataset runs through December 2011, we do not have 36 months of synthetic or actual data in the post-intervention period because our outcome variable is a twelve-month centered-moving average of sales volume.

We are interested in understanding not just the direction of the effect of Chrysler's acceptance of government assistance, but also the magnitude of the effect. In order to better understand how large of an impact on sales Chrysler's acceptance of government funds had, we examine the size of the gap between the actual firm's sales and the synthetic firm's sales in Figure 3. At the gap's widest, Chrysler sold 40,000 fewer vehicles in that month than the counterfactual synthetic suggests it would have had the firm forgone government assistance. Given the level of Chrysler's sales prior to the intervention, Chrysler appears to have sold 20%fewer vehicles, or worse, relative to its counterfactual in each month following its acceptance of government assistance. This gap is large—and has significant implications for management researchers' understanding of the dynamic between private firms and public institutions more broadly.

Any concerns with premature divergence between the synthetic and the actual Chrysler, in Figures 2 or 3, can be alleviated (i) by recognizing that there is a similarly sized positive deviation between the actual and synthetic earlier in the pre-intervention period; and, (ii) by recalling that to smooth out seasonal and promotional fluctuations in the sales data we used a twelve month centered moving average—meaning that at a given date some future performance is included.

<Insert Figure 2 Here>

<Insert Figure 3 Here>

3.3.2 Robustness

Robustness checks are particularly important to conduct when using synthetic control since we are making causal inferences based on subsamples of small *n* datasets. To verify the causal effect of the treatment in synthetic control we can run placebo tests—which check whether or not the same results would manifest had the intervention occurred at a different point in time or among untreated statistical units. These placebos should not respond to false interventions in the same way that the focal unit does to the actual event if the effect is genuine. Further robustness checks can falsify other underlying assumptions.

Placebo Tests in Time

One way to check for robustness of results is to introduce a false intervention at a different point in time. In this case, there should be no effect of the false timing of Chrysler accepting government assistance if the timing of the effect is causal.[17]

[17] Abadie, et al. (2011) do this with respect to the date in which Germany was re-unified in an analysis of its effects on West German GDP per capita, showing that the was no divergence around "hypothetical" re-unifications a decade or two earlier.

In Figure 4, we treat the intervention as if it had occurred 4 years (or 48 months) earlier in time, when we know it did not occur then in reality. To implement this placebo in-time test, we simply re-run our analysis using January 2005 as the intervention date instead of January 2009, while continuing to use a 48 month window prior to the intervention date on which to minimize prediction error.

Figure 4 clearly shows that the placebo intervention in January 2005 does not have a dramatic effect on Chrysler's future sales since the synthetic version of the firm continues to track the performance of the actual firm quite closely. The absence of real divergence improves our confidence in our core result: Chrysler's acceptance of government assistance in January 2009 caused the firm's sales to deteriorate by more than they would have if the firm had not accepted it at that date.

<Insert Figure 4 Here>

Placebo Tests among Untreated Statistical Units

Another way to conduct a robustness check is to run a placebo test among untreated statistical units. The purpose is to check if the intervention had been falsely applied to each of the control units, rather than the focal unit, whether there would be as discernibly negative a divergence between the synthetics for the control units and the synthetic for the focal unit. If the effect of the treatment on the focal unit is causal, then we would not expect an application of the treatment to untreated units to lead to an equally large divergence (or treatment effect).[18]

Figure 5 illustrates the gap in performance between the actual and the synthetic versions of each of the sixteen possible control firms. In it, synthetics for each control firm are created as if they had received an intervention in January 2009 just as Chrysler did in reality. We continue to use a 48 month pre-treatment window over which to minimize pre-treatment prediction error. The solid line represents the gap in performance between the actual and synthetic Chrysler as depicted in Figure 3. Figure 5 demonstrates that the gap in Chrysler's performance relative to its synthetic's following the January 2009 acceptance of government assistance is far greater than the gap for any of the control units to which a placebo intervention was applied. This further bolsters our causal interpretation of the effect of Chrysler accepting conditional government assistance as having a more significant detriment on their sales than at control firms, which received hypothetical treatments.

<Insert Figure 5 Here>

[18] Abadie et al. (2010) do this with respect to the state in which anti-tobacco legislation was passed in an analysis of the impact of a California law's effects on cigarette consumption by hypothetically assuming every other state passed a similar law at the same time.

In Figure 6, we present the results of the placebo test, among untreated firms, excluding synthetics for control firms with mean-squared prediction errors (MSPE) in the pre-intervention period greater than 20 times that of Chrysler—because several control firms in Figure 5 have a poor synthetic fit.[19] This filter removes noisy observations to clarify the result that placebo interventions among untreated units do not have as large of a negative effect on sales.[20]

<Insert Figure 6 Here>

The data generated in running the placebo-among-units test also provides a basis for evaluating the strength of inferences quantitatively. In a traditional regression setting, we assess the strength of inferences with p-values that capture the fraction of observations under the distribution of all possible observations that are at least as extreme as a specified data point; we can do the same in synthetic control—generating p-values that assess the probability of obtaining a result at least as extreme as the one identified for the focal unit in the event that the treatment were randomly assigned to any observation unit in the population.[21]

To evaluate the relative extremity of the treatment event on each observation unit, we first calculate a scale-independent measure reflecting treatment extremity so we can compare observation units directly with each other. We use the ratio of the root mean square prediction error (RMSPE) for the post-treatment period to the RMSPE for the pre-treatment period. We construct this measure for each firm in the population evaluated in the placebo-among-firms test above. We may also want this extremity measure to take into account whether the underlying treatment effect was positive or negative on average, since positive extremes may be interpreted very differently from negative extremes; multiplying the ratios of post-treatment/pre-treatment RMSPEs by -1 for observation units where the mean treatment effect is negative brings this directional information back into the measure. We can interpret values of this measure close to 1 or -1 as indicating there are no major changes in how well the synthetic matches the actual unit in the post-period relative to the pre-period.

This treatment extremity data provides the basis for constructing a probability distribution of the extremity of post-treatment event outcomes for each observation unit, or firm in our case. Since we only have a discrete number of outcomes, this probability distribution can be represented as a histogram rather than a continuous distribution. In Figure 7, we display the histogram of our

[19] Poor fits in the pre-prediction period can be expected in synthetic control analysis, particularly for observation units at the extremes (Abadie et al. 2010).

[20] The filter for exclusions that we apply here at 20 times MSPE is conservative; Abadie et al. (2010), in their study of state-level anti-tobacco laws, exclude control units with MSPE that are only 2 times greater than California which was the focal unit.

[21] Abadie et al. (2010) are the first to calculate such a metric, but do not explicitly referred to it as a p-value until Abadie et al. (2011). Those authors note that the technique is based on Rosenbaum's (2002a, 2002b) prescribed methods for permutation inference in either randomized experiments or observational studies.

treatment extremity measure for the 17 firms in our population, 16 of which are control firms and 1 of which is the focal firm, Chrysler. The majority of the observations fall in the range [-3,3] indicating that the government assistance program had negligible effects on most auto manufacturers. Chrysler, represented in black, occupies the second most negative position with a treatment extremity value of -5.0.

Recognizing that there are 17 firms in the population and that Chrysler is second from the bottom of the distribution, we can readily calculate our p-value, as we would in a regression setting, as 2/17 ≈ 0.12. Hence, the probability of seeing as large of a negative effect of government assistance on sales as we saw at Chrysler is approximately 12% if instead the government assistance were randomly assigned to any firm in the population. While in traditional regression settings many researchers would reject null hypotheses when p-values are greater than a specified level (typically p>0.05), we need to be particularly careful in synthetic control settings to calibrate our interpretation of p-values (Sellke et al. 2001). The negative effect of government assistance on Chrysler's sales can still be interpreted as being measurably greater than zero in this instance with a p-value of 0.12.[22]

<Insert Figure 7 Here>

Further Robustness Checks

Other potential robustness checks aimed at falsifying underlying assumptions include : (i) leave-one-out tests; and (ii) out-of-sample validation (Abadie et al. 2011). Leave-one-out tests check whether the synthetic is particularly sensitive to the inclusion of a specific control unit (or outlier) in its construction; this may be a particular concern if one of the units that comprise the synthetic experienced a large, unrelated shock in the post-treatment period.[23] Out-of-sample validation takes data from well before the treatment is applied, leaving a gap before the event, and uses this data to construct a synthetic; this helps alleviate any concerns about the period immediately preceding the

[22] We need to be particularly careful not to make Type II errors by falsely rejecting real treatment effects by calibrating our interpretation of p-values in this synthetic control setting. To illustrate why we need to avoid the p-value fallacy here, consider the hypothetical case where Chrysler had the most negatively extreme relative response to the government assistance applied in January 2009; it's p-value would still only be 1/17 ≈ 0.06 in that scenario. If we followed a p>0.05 rule we would reject the null hypothesis and declare the treatment effect to be negligible despite it being the single most extreme post-treatment realization. Moreover, in our data the most-negatively extreme observation unit in the post-January 2009 period was BMW—a firm whose underperformance in this timeframe has been attributed to factors other than government assistance, namely the competitive resurgence of Volkswagen-Audi who took market share from BMW in the luxury segment (Libby 2012). Unsurprisingly, Volkswagen-Audi occupies one of the most extremely positive positions in Figure 7.

[23] An implicit assumption in conducting synthetic control is that any shocks in the post-treatment period are either (i) common to all units including the treated or (ii) minor and offsetting among control units. Leave-one-out tests explicitly check whether or not these assumptions hold. We may worry the assumptions do not hold if a synthetic constructed leaving-one-out were dramatically different from the synthetic constructed without leaving any out.

event being abnormal and about whether or not there are omitted, time-varying concomitant-variables.[24] These further robustness checks support our findings when we implement them.[25]

4 DISCUSSION

Our analysis provides the first robust evidence demonstrating that Chrysler would have performed better had it declined to accept government assistance in January 2009: the firm would have sold approximately 20% more vehicles in each of the first thirty months following its acceptance of government assistance had it made the alternative choice. Another way to interpret the synthetic's counterfactual sales data are not necessarily in thinking about them as representing the sales produced by a legal entity named Chrysler, but rather produced using the firm's constituent parts, i.e. production platforms, brands, technology, etc., as of the intervention date. Hence, we would expect the synthetic Chrysler—based on the actual Chrysler's pre-assistance constituent parts—to generate the same sales volume whether Chrysler stayed intact as a firm (had it been able to find sufficient private sector financing) or disappeared altogether (with components sold to other auto manufacturers or shutdown entirely). In fact, in the post-treatment period, one control unit that comprises part of the synthetic Chrysler, Isuzu, did shutdown their North American operations, while other control units that comprise the synthetic Chrysler thrived under different managerial conditions (i.e. without government involvement) at other auto manufacturers.

The insights the synthetic control analysis provides stop at quantification of the magnitude and direction of the effect of government assistance on Chrysler, leaving other questions unanswered. What were the underlying causal mechanisms that drove this result? Can we say anything about the effect of government assistance on General Motors' sales? And, can we say whether it was a good idea from the government's perspective to extend the assistance packages in the first place? While we can speculate on answers, complementary analysis using other methods with different strengths and applying them to different data sets may better help resolve these questions.

4.1 Exploring Mechanisms underlying Declining Sales at Chrysler

While synthetic control can be used to estimate, with relative precision, the magnitude and direction of a treatment effect, it does not directly reveal underlying causal mechanisms. Hence, building theory on causal processes from synthetic control results remains an outstanding challenge; detailed qualitative analysis is a complementary approach that could rise to that challenge. Building

---

[24] We might worry about either of these concerns if the synthetic constructed using out-of-sample data did not approximate the synthetic using in-sample data.

[25] All robustness checks will be included in the online supplemental materials containing data and STATA .do files that allow interested readers to replicate our entire analysis.

on the results from our synthetic control analysis, qualitative researchers would benefit from focusing on the four control firms with positive weights in the synthetic Chrysler. Examining this sub-set of auto manufacturers and what they did in the post-intervention period could be fruitful—because we can think of these firms as if they were separate divisions of the counterfactual Chrysler whose relative performance we constructed. This approach could help uncover the specific organizational nuances that were missing at Chrysler.[26] We could then ground theory regarding the mechanisms in the differences between organizational nuances at Chrysler vis-à-vis its synthetic components, which we would be treating as sub-divisions of the synthetic Chrysler.

Without applying a qualitative approach to develop grounded theory, we can nevertheless use existing theories to build explanations for the causal effect we identified. Potential mechanisms explaining why Chrysler sold fewer vehicles because it took government assistance include: consumer backlash, governance and compensation challenges, lack of consensus among key decision makers, and competitive reactions to government involvement.

The social movement literature predicts resistance to grow against firms when issues they are associated with become salient and sets of actors can unite around shared beliefs, which are often political in nature (Davis & Thompson 1994). That prediction is consistent with one argument for Chrysler's underperformance: that consumer backlash against the firm for taking a tax-payer funded "bailout" weighed on sales despite those who believed government backing would assuage consumer's fears about warranty and maintenance issues (Stewart 2012). Indeed, independent market research found that 22% of likely purchasers planned to avoid Chrysler products because of the firm's decision to accept government assistance (Vlasic 2012).[27]

Corporate governance scholars predict that firm performance is a function of executive compensation and related incentive structures (Jensen & Murphy 1990). Hence, any exogenous restrictions imposed on executive compensation could significantly impact performance if attracting and retaining top managerial talent is critical (Pfeffer 1994). These mechanisms explaining firm performance are consistent with an argument that Chrysler underperforms because government conditions on executive compensation were overly restrictive, making it difficult for the firm to remain competitive in the market for executive talent. On October 10, 2009, the Treasury's Special Pay Master announced rules capping cash salaries at $500,000 for most executives at the seven biggest recipients of TARP funds (including Chrysler). Within five months there was over 15%

[26] Qualitative researchers would want to focus on nuances relating to attributes receiving high weights in v_m produced using synthetic control—as these attributes are quantitatively the best predictors of outcome measures.

[27] This finding comes from a survey question included on a broader, ongoing automotive industry market research questionnaire administered in the first quarter of 2012 by CNW Marketing Research of Brandon, Oregon.

turnover among top management at these firms (Treasury Press Center 2010), a rate that is over five times higher than typical at large firms.[28]

Greater consensus among top decision makers contributes to superior performance (Priem 1990), whereas politicking among decision makers with divergent interests often leads to poor performance (Eisenhardt & Bourgeios 1988). Including the Treasury in the set of decision makers at Chrysler could therefore have impeded performance if government and private sector shareholders pursued different objectives. Tensions between the Treasury and Chrysler management occurred over several strategic and operational issues. One example of a decision making conflict involved the firm's 255,000 pension obligations that were underfunded by $9.3B. Management initially wanted the pension obligations discharged in bankruptcy proceedings because this would maximize flexibility and profitability going forward;[29] however, the government wanted to avoid that outcome, because the federal Pension Benefit Guaranty Corporation (PBGC) would have been responsible for covering $2B of the pension obligation shortfall, leaving the remaining $7.3B burden on union and former-union employees (Walsh 2009; Whoriskey 2009). By the time Chrysler emerged from managed bankruptcy proceedings, it had made concessions: a United Auto Workers' Retiree Fund was granted an initial 67.69% equity stake in the reorganized firm as compensation for the firm's unfunded obligations thereby protecting the PBGC at the cost of the firm (Anginer & Warburton 2010; Roe & Skeel 2010). A second issue that involved a dispute between the government and Chrysler management was over the plan to rationalize an over-extended dealer network. The Treasury determined that Chrysler's proposal in early 2009 for closing dealerships was too slow, requiring Chrysler to accelerate its plans by immediately terminating 789 dealerships, rather than gradually phasing them out over a five year period (SIGTARP 2010). Chrysler executives disputed the Treasury's position, warning that closing dealerships too rapidly would lead to a loss of market share particularly in rural areas, a relative stronghold (SIGTARP 2010). The Treasury's requirement that management follow its direction on the pace of dealership closures ultimately led to state-level legal challenges for the firm, diverting managerial attention from operational performance (Black 2010).

Seamans (2012) finds that the emergence of government-owned, or financed competitors, can alter the competitive landscape by causing other firms to compete more vigorously. This may help explain Chrysler's sales shortfall following government intervention. Competitive reactions to the

[28] Wiersama and Bantel (1993) report an average turnover rate of 20% for members of top management teams at Fortune 500 firms over a three year window, equivalent to a 2.8% rate over a five month window assuming turnover is uniform across time.
[29] Similar arguments for discharging pension liabilities in had been made successfully in prior bankruptcy proceedings by the likes of United Airlines in 2005 and Bethlehem Steel in 2002.

government assistance programs may account for why some firms that did not accept government assistance or were not offered government assistance outperformed their synthetics in the post-government intervention period, as shown in Figures 5 and 6. Hence, in a competitive environment, the treatment is applied not to one observation unit, Chrysler, but to all observation units, the entire auto industry. Ford responded to the government intervention by focusing its own vehicle redesign efforts specifically on segments in which its government-assisted rivals had recently gained market share (Terlep & Ramsey 2011). Moreover, Ford ran a marketing campaign positioning themselves as the only American firm that did not take government assistance.[30]

4.2 Extending the Analysis to General Motors

Given our finding that Chrysler, or its surviving components, would have sold 20% more vehicles had they forgone government assistance and our exploration of potential mechanisms, it is natural to wonder what the result would have been for General Motors, as they also took government assistance, and whether the same mechanisms would apply. Unfortunately, the General Motors case is one where the synthetic control method cannot be applied fruitfully. Recall that the method cannot be used to analyze extreme statistical units—and note that General Motors is such a unit as it is 50% larger than any other auto manufacturer. The reason why synthetic control comes up short in these extreme cases is that any weighted average of control units, where the weights must sum to one, can never achieve the same scale as the extreme statistical units. Attempting to create a synthetic in this situation yields one with unacceptable pre-intervention prediction error. Abide et al. (2010) note this problem with respect to New Hampshire in their smoking study focused on California. While we can't apply synthetic control effectively to the General Motors' case, our synthetic control analysis of Chrysler may still inform a qualitative study of that firm.

4.3 Extrapolating to the Public Policy Debate

While we might be tempted to stretch our finding beyond the managerial question we asked, we are limited in how broadly we can interpret our results. We cannot fully resolve the hotly-debated public policy question: whether or not the government made the right decision in offering assistance to Chrysler. We use synthetic control to analyze how Chrysler's sales would have evolved had the firm not taken government assistance. The government may, however, have had objectives

[30] In one of Ford's "Drive One" television advertisements featuring real customers who are put on the spot about their decision to buy from the company, a Ford F-150 owner named Chris explains "I wasn't going to buy another car that was bailed out by our government. I was going to buy from a manufacturer that's standing on their own: win, lose, or draw. That's what America is about is taking the chance to succeed and understanding when you fail that you gotta' pick yourself up and go back to work. Ford is that company for me." (Bedard 2011)

in addition to maximizing the firms' sales; namely, policymakers may have been interested in keeping an iconic American firm in business, ensuring the economic viability of Michigan's economy, protecting existing union employment, and improving re-election prospects. What our analysis can tell us about the policy debate is that if selling more vehicles would have helped achieve policy goals, then the assistance program may have fallen short in this one aspect.

5 CONCLUSION

In this paper, we introduced the synthetic control method to management research in order to answer a question existing methodologies often cannot answer: what impact does government involvement have on firm performance? Specifically, we estimate that Chrysler, or its remaining constituent parts, would have sold 20% more vehicles had the firm not taken conditional government assistance in the midst of the recent financial crisis. In addition, we provided direction for how to apply synthetic control methodology more broadly. In doing so, we demonstrated what concerns researchers should consider in robustness checks when applying the method. We have made our data and code publicly available to enable other researchers to learn how to implement the method and how to check the robustness of their results.

The synthetic control method fills an important gap in that it enables management researchers to better analyze the effects of phenomena that occur in limited populations and/or apply to only a small number of observation units. Previously management researchers had limited tools to accurately assess the magnitude and direction of such phenomena's effects. Nevertheless, this is a particularly important task in empirical settings when competing theories predict big/small or positive/negative effects as was true in our context.

While we applied synthetic control in a single context, examining a firm and its interaction with its external environment, many other applications remain to be explored in alternative contexts. We considered the firm the unit of analysis; however, the method can also be applied at an industry level, within divisions of firms, or even among individuals employed at firms. The treatment in our example was externally determined, although other treatments need not be. The method could be applied to: industry self-regulation; technological breakthroughs at a firm or the adoption of new practices (e.g. ISO 9001); outsourcing decisions made differentially across divisions or facilities; or, modification of individuals' incentive schemes. It may also be useful to address the growing interest in rare events (Lampel et al. 2009; Rerup 2009) given other approaches' methodological limitations. Finally, there may be cost-saving practitioner applications to synthetic control where managers can explore new tactics on a sub-set of the firm and measure the associated cost or benefits.

6 REFERENCES

Abadie, A., A. Diamond, and J. Hainmueller. 2010. “Synthetic Control Methods for Comparative Case Studies: Estimating the Efect of California's Tobacco Control Program." *Journal of the American Statistical Association*. 105(490): 493-505.

Abadie, A., A. Diamond, and J. Hainmueller. 2011. “Comparative Politics and the Synthetic Control Method” *MIT Department of Political Science Working Paper No. 2011-25.* Available online at: http://ssrn.com/abstract=1950298

Abadie, A. and J. Gardeazabal. 2003. “The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review*. 93(1): 112-132.

Anginer, D., A. J. Warburton. 2010. "The Chrysler Effect: The Impact of Government Intervention on Borrowing Costs." Working Paper. [Available online at: http://ssrn.com/abstract=1640801]

Bedard, P. 2011. “Ford TV Ad Slams Obama Auto Bailouts.” *US News & World Report’s Washington Whispers Blog:* 16 September 2011. [Available online at: http://www.usnews.com/news/blogs/washington-whispers/2011/09/16/ford-tv-ad-slams-obama-auto-bailouts]

Black, B. 2010. “The U.S. as “Reluctant Shareholder”: Government, Business and the Law”. *Entrepreneurial Business Law Journal*, 5(2): 561-593

Davis, G.F., T.A. Thompson. 1994. “A Social Movement Perspective on Corporate Control.” *Administrative Science Quarterly.* 39(1): 141-173.

Dolan, M. 2009. “Ford Targets Rivals as Loss Eases.” *Wall Street Journal*. 25 April 2009.

Donohue III, J. J. and Aneja, A. 2011. “New Evidence in the ‘More Guns, Less Crime’ Debate: A Synthetic Controls Approach” Working Paper: Stanford Law School. [Available on SSRN at: http://ssrn.com/abstract=1984969]

Eisenhardt, K. M. 1989. “Building Theories from Case Study Research.” *Academy of Management Review.* 14(4): 532-550.

Eisenhardt, K. M. 1988. “Politics of Strategic Decision Making in High-Velocity Environment: Toward a Midrange Theroy.” *Academy of Management Journal*. 31(4): 737-770.

Eisenhardt, K. M., and Graebner, M. E. 2007. “Theory Building from Cases: Opportunities and Challenges.” *Academy of Management Journal.* 50(1): 25-32.

Fiss, P. 2007. “A Set Theoretic Approach to Organizational Configurations.” *Academy of Management Review.* 32(4), 1180–1198.

Heckman, J.J. and V.J. Hotz. 1989. “Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training." *Journal of the American Statistical Association.* 84(408): 862-874.

Jensen, M. C., K. J. Murphy. 1990. “Performance Pay and Top-Management Incentives.” *Journal of Political Economy*. 98(2): 225-264.

Kivleniece, I., B.V. Quelin. 2012. “Creating and Capturing Value in Public-Private Ties: A Private Actor’s Perspective.” *Academy of Management Review.* 37(2): 272-299.

Lampel, J., J. Shamsie, Z. Shapira. 2009. "Experiencing the Improbable: Rare Events and Organizational Learning." *Organization Science.* 20(5): 835-845.

Lang, L. H. P., R. Stulz. 1994. "Tobin's q, Corporate Diversification, and Firm Performance." *Journal of Political Economy.* 102(6): 1248-1280.

Laeven, L., R. Levine. 2007. "Is there a diversification discount in financial conglomerates?" *Journal of Financial Economics.* 85(2): 331-367.

Libby, Tom. 2012. "Audi versus BMW, Mercedes-Benz." R.L. Polk & Co. Automotive Knowledge Blog: 21 March 2012. [Available online at: http://blog.polk.com/blog/blog-posts-by-tom-libby/audi-versus-bmw-mercedes-benz]

Mahoney, J., A. McGahan, C. Pitelis. 2009. "The Interdependence of Private and Public Interests." *Organization Science.* 20(6): 1034-1052.

March, J.G., L.S. Sproull, and M. Tamuz. 1991. "Learning from Samples of One of Fewer." *Organization Science.* (2): 1, 1-13.

Marx, M., D. Strumsky, L. Flemming. 2009. "Mobility, Skills, and the Michigan Non-Compete Experiment." *Management Science*. 55(6): 875-889.

Office of the Press Secretary of the White House. 2008. *Fact Sheet: Financing Assistance to Facilitate the Restructuring of Auto Manufacturers to Attain Financial Viability*. Released: 19 December 2008 [Available in online archives at: http://georgewbush-whitehouse.archives.gov/news/releases/2008/12/20081219-6.html]

Pfeffer, J. 1994. *Competitive Advantage through People.* Harvard Business School Press: Boston.

Priem, R. L. 1990. "Top Management Team Group Factors, Consensus, and Firm Performance." *Strategic Management Journal.* 11(6): 469-478.

Rerup, C. 2009. "Attention Triangulation: Learning from Unexpected Rare Crises." *Organization Science.* 20(5): 876-893.

Roe, M. J., D. Skeel. 2010 "Assessing the Chrysler Bankruptcy." *Michigan Law Review* 108: 721-771.

Romney, M. 2012. "Romney op-ed: U.S. autos bailout 'was crony capitalism on a grand scale.'" *The Detroit News: 14 February 2012.*

Rosenbaum, P. R. 2002a. *Observational Studies* (2nd ed.) New York: Springer.

Rosenbaum, P. R. 2002b. "Covariance Adjustment in Randomized Experiments and Observational Studies," *Statistical Science.* 17 (3): 286–327.

Runde, J. and de Rond, M. (2010) "Evaluating causal explanations of specific events," *Organization Studies* 31(4): 431-450

Seamans, R. 2012. "Fighting City Hall: Entry Deterrence and Technology Deployment in the Cable TV Industry," *Management Science* 58(3) 461-475.

Sellke, Thomas; Bayarri, M.J.; Berger, James (2001). "Calibration of p values for testing precise null hypotheses". *The American Statistician.* 55 (1): 62–71.

SIGTARP (Office of the Special Inspector General for the Troubled Asset Relief Program). 2010. Factors Affecting the Decisions of General Motors and Chrysler to Reduce their Dealership Networks. SIGTARP-10-008: 19 July 2010. [Available online at:

http://www.sigtarp.gov/Audit%20Reports/Factors%20Affecting%20the%20Decisions%20of%20General%20Motors%20and%20Chrysler%20to%20Reduce%20Their%20Dealership%20Networks%207_19_2010.pdf]

Sullum, J., M. Welch, R. Bailey. "Illegal. Illiberal. Ill-Fated. Why Washington shouldn't run Detroit." *Reason*. 41(4): 20

Stewart, J. B. 2012. "When Debating the Auto Bailout, Consider Lehman's Fate." *New York Times:* 10 March 2012.

Terlep, S. and M. Ramsey. 2011. "Ford and GM Renew a Bitter Rivalry." *Wall Street Journal:* 23 November 2011.

Thorngate, W. 1976. "'In General' vs. 'It Depends': Some Comments of the Gergen-Schlenker Debate." *Personality and Social Psychology Bulletin*. 2: 404-410.

Treasury Press Center. 2010. "Special Master Issues 2010 Rulings for 'Top 25' Executives at Firms Receiving Exceptional Taxpayer Assistance and 'Look Back' Letter on Review of Pre-Recovery Act Compensation." Released March 23, 2010. [Available in online archives at: http://www.treasury.gov/press-center/press-releases/Pages/tg604.aspx]

Villalonga, B. 2004. "Diversification discount or premium? New evidence from the business information tracking series." *Journal of Finance*. 59: 479–506

Vlasic, Bill. 2012. "A Weight Hobbling G.M." *New York Times*: 3 May 2012.

Walsh, M.W. 2009. "Plight of Carmakers Could Upset All Pension Plans." *New York Times*: 24 April 2009.

Weick, K. 1979. *The Social Psychology of Organizing*. Addison-Wesley: Reading, MA.

Wiersema, M.F., K.A. Bantel. 1993. "Top Management Team Turnover as an Adaption Mechanism: The Role of the Environment." *Strategic Management Journal* 14: 485-504.

Whoriskey, P. 2009. "Auto Retirees Brace for Hardship; Pension Benefits May Suffer in Bankruptcy." *Washington Post*. 24 April 2009.

7 APPENDIX: DATA DESCRIPTION AND SOURCES

Variable	**Description**	**Source**
Light vehicle sales volume	Total units sold by firm of new vehicles in the U.S., monthly	*Ward's Auto: U.S. Model Car Specifications and Prices*
Average price	Average retail price of all vehicles sold (U.S. dollars), monthly	Author's calculations based on data from *Ward's Auto: U.S. Model Car Specifications and Prices*
Average fuel economy	Average fuel economy of all vehicles sold in the U.S. (miles per gallon), monthly	Author's calculations based on data from *Ward's Auto: U.S. Model Car Specifications and Prices*
Maximum fuel economy	Maximum fuel economy of all vehicles sold in the U.S. (miles per gallon), monthly	Author's calculations based on data from *Ward's Auto: U.S. Model Car Specifications and Prices*
Average weight	Average weight of cars sold in the U.S. (lbs.), monthly	Author's calculations based on data from *Ward's Auto: U.S. Model Car Specifications and Prices*
Average engine size	Average size of engines of all vehicles sold in the U.S. (Liters), monthly	Author's calculations based on data from *Ward's Auto: U.S. Model Car Specifications and Prices*
Number of active production platforms	Total number of production platforms for vehicles sold in the U.S., monthly	Author's calculations based on data from *Ward's Auto: U.S. Model Car Specifications and Prices*
Number of active brands	Total number of brands marketed by firm in the U.S. (e.g. In 2011, Chrysler markets the Chrysler, Dodge, Jeep, Ram and Fiat branded vehicles), monthly	Author's calculations based on data from *Ward's Auto: U.S. Model Car Specifications and Prices*
Number of active series	Total number of series marketed within a brand by a firm in the U.S. (e.g. In 2011, Chrysler branded vehicles include four series: 200, 300, Sebring and the Town and Country.), monthly	Author's calculations based on data from *Ward's Auto: U.S. Model Car Specifications and Prices*
Number of market segments	Total number of market segments that a firm competes within in the U.S. (e.g. In 2011, Chrysler competed within 14 different segments, including small specialty, middle specialty, upper small, upper middle, small SUV, small CUV, small van, middle SUV, middle CUV, small pickup, medium duty, large SUV, large regular, and large pickup), monthly	Author's calculations based on data from *Ward's Auto: U.S. Model Car Specifications and Prices*
Fraction of vehicles sold from the SUV, light truck, and van segments	Proportion of total U.S. sales that are made in the larger vehicle segments, monthly	Author's calculations based on data from *Ward's Auto: U.S. Model Car Specifications and Prices*
Fraction of vehicles manufactured outside North America	Proportion of total U.S. sales of imported vehicles, monthly	Author's calculations based on data from *Ward's Auto: U.S. Model Car Specifications and Prices*
Number of employees	Total number of worldwide employees, annual	*Compustat*
Leverage ratio	Total Long-term debt / Total Assets, annual	*Compustat*

TABLES AND FIGURES

Table 1: Comparison of Attributes between Chrysler, Ford, and Synthetic Chrysler in the period prior to Acceptance of Government Assistance

	(Averages of Monthly Figures over 48 months prior to acceptance of government assistance)		
Attribute	*Chrysler*	*Synthetic Chrysler*	*Ford*
Price, Average of Vehicles Sold	26843	27368	28384
MPG, Average of Vehicles Sold	19.9	20.3	19.4
MPG, Maximum of Vehicles Sold	31.5	34.8	34.6
Weight (in lbs), Avg. of Vehicles Sold	4181	4194	4422
Engine Size (in L), Avg. of Vehicles Sold	3.95	3.96	4.35
Fraction of Sales in SUV/Truck/Van	0.751	0.676	0.657
Brands, # Active	4.3	2.6	3.0
Platforms, # Active	15.8	15.3	18.1
Segments of Market, # Active	15.3	14.0	16.0
Series of Vehicles, # Active	22.3	20.7	24.3
Fraction of Sales Manufactured Outside North America	0.008	0.094	0.000
Leverage (Debt/Assests)	0.249	0.323	0.391
# of Employees	277211	259439	323050
Sales Volume, Moving Avg. (12 Month)	180582	194465	251474
Dec '08 Level of Sales Volume, Moving Avg. (12 Month)	149.7	240.4	161.5

Table 2: Weights of Companies in Synthetic Chrysler

Company	*Weight*	*Company*	*Weight*
BMW	0	Mitsubishi	0
Daimler	0	Nissan	0
Ford	0.655	Porsche	0
General Motors	-	Saab	0
Honda	0	Subaru	0
Hyundai-Kia	0	Suzuki	0.021
Isuzu	0.162	Toyota	0.152
Jaguar Land Rover	-	Volkswagen-Audi	0
Mazda	0	Volvo	0

Notes:
(1) General Motors is excluded as a match because it also received government assistance
(2) Jaguar Land Rover is excluded as match because of limited data availability

Figure 1: Chrysler and Ford Sales around Chrysler's Acceptance of Government Assistance

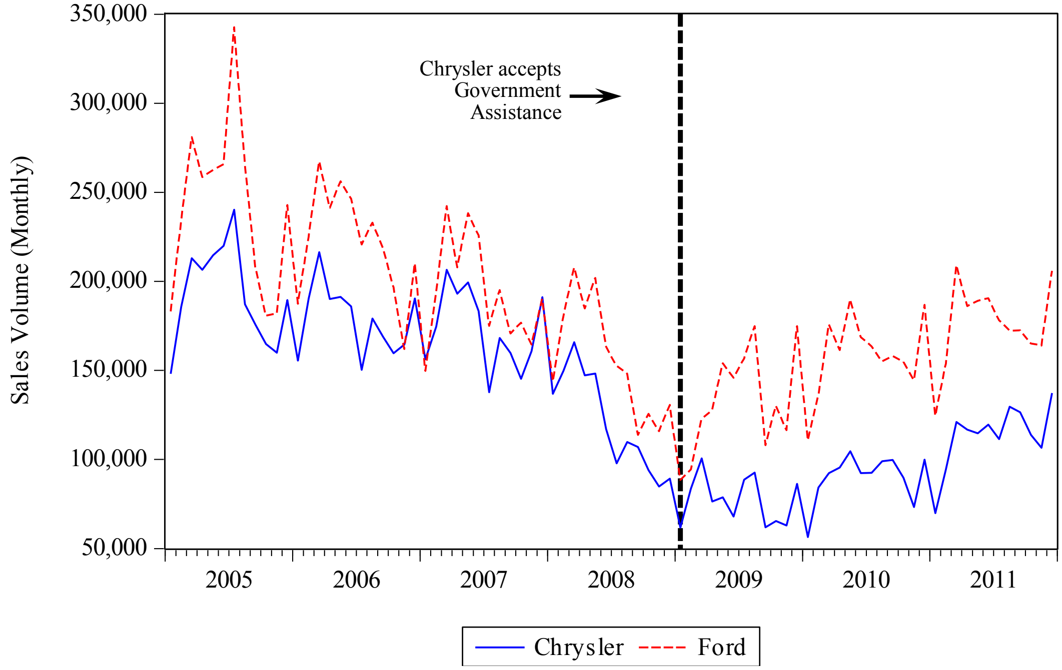

Figure 2: Chrysler's & Synthetic's Sales Volume around Acceptance of Government Assistance

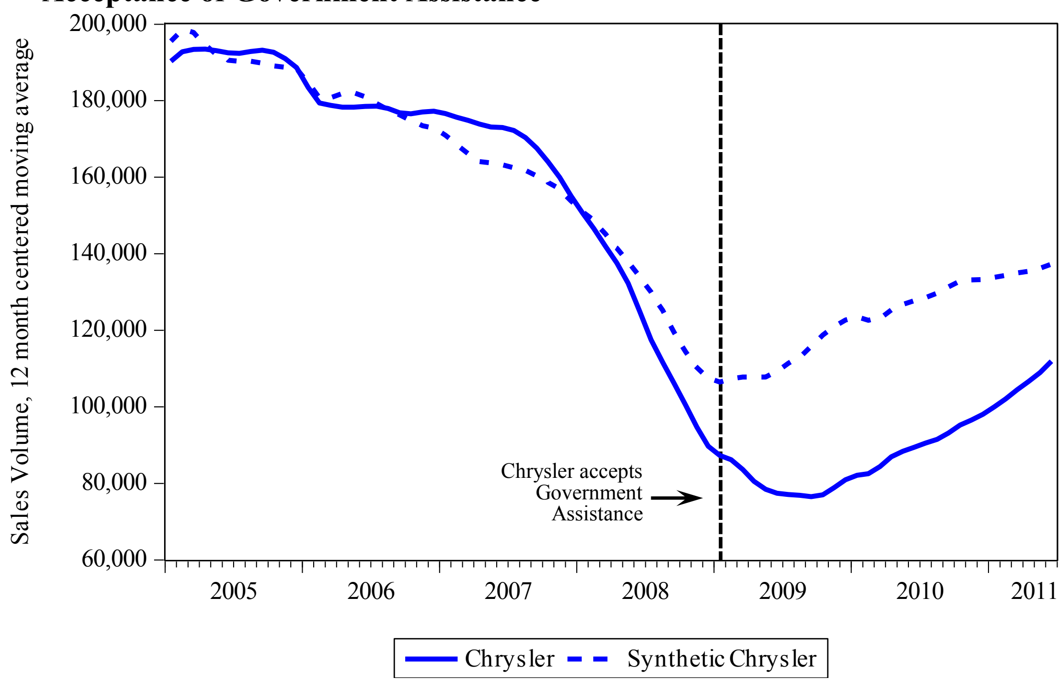

Figure 3: Difference between Chrysler's Actual & Synthetic Sales Volume around Acceptance of Government Assistance

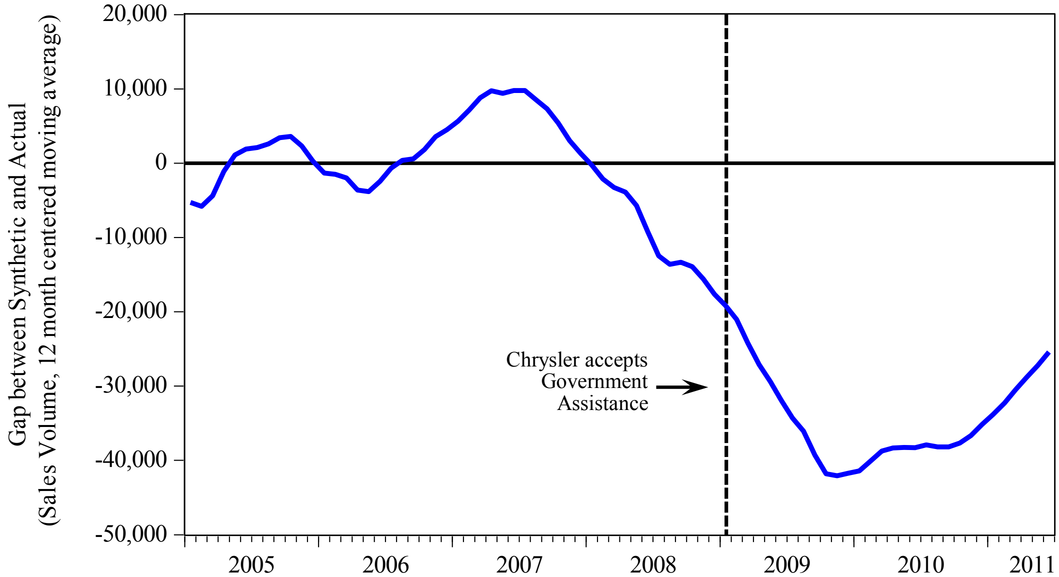

Figure 4: Placebo Test, In Time. Chrysler & Synthetic's Sales Volume around *Hypothetical* January 2005 Acceptance of Government Assistance

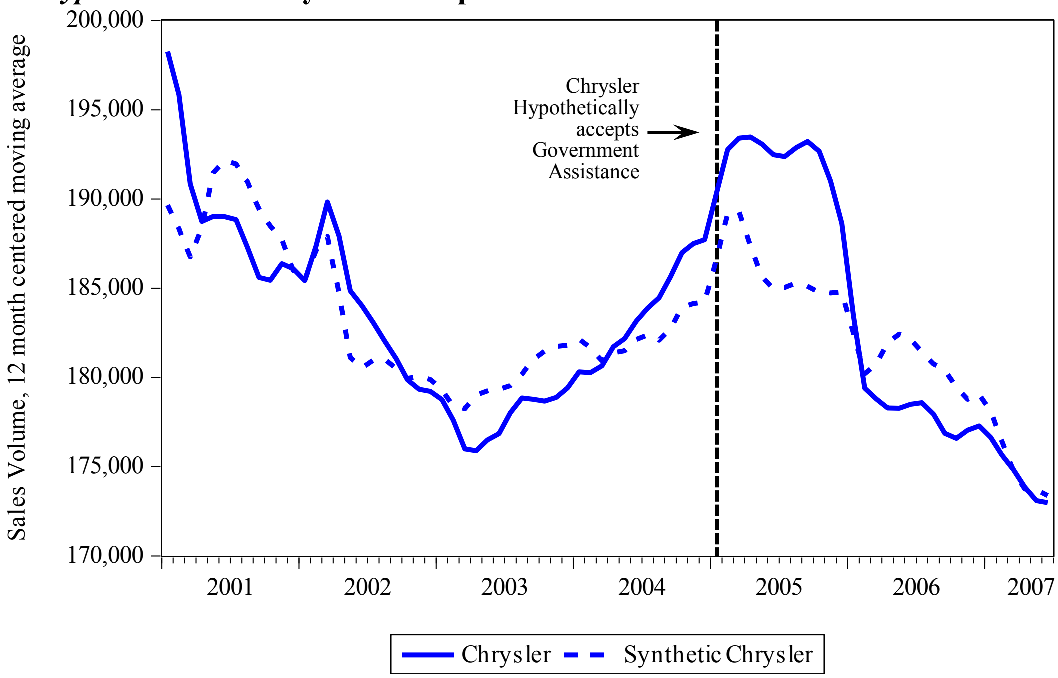

Figure 5: Placebo Test, Among Untreated Firms. Difference between Actual & Synthetic Sales Volume around January 2009 Acceptance of Government Assistance

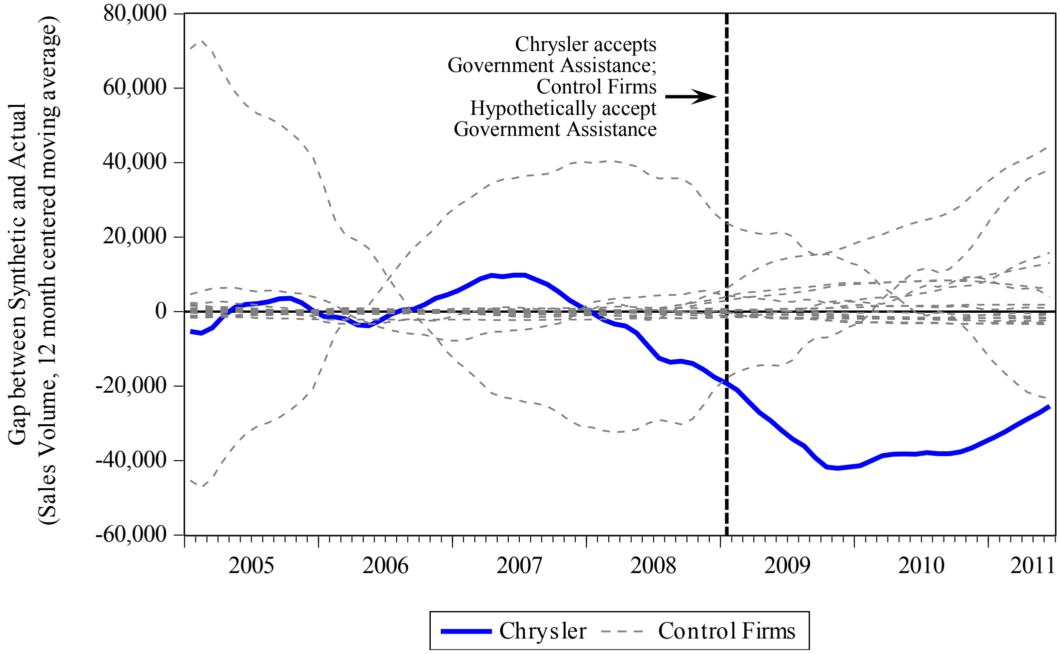

Figure 6: Placebo Test, Among Untreated Firms. Difference between Actual & Synthetic Sales Volume around January 2009 Acceptance of Government Assistance (Excluding Firms with MSPE 20x higher than Chrysler's prior to Acceptance of Government Assistance)

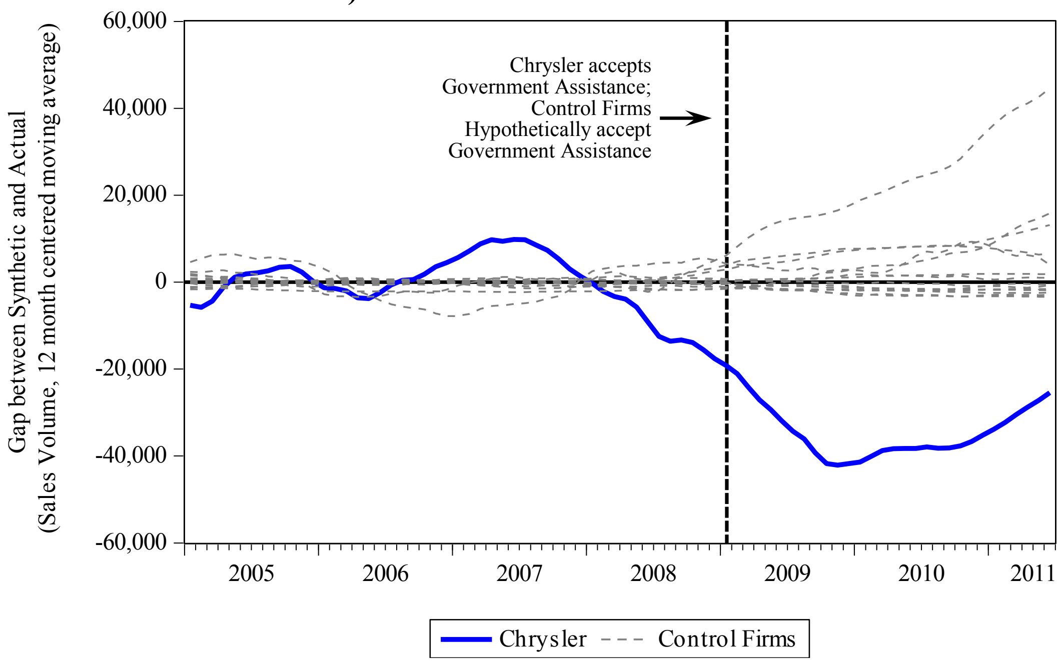

Figure 7: Histogram of Treatment Extremity Measure, Among-Firm Placebos' Post-Treatment Performance Relative to Pre-Treatment Performance