Available online at www.sciencedirect.com

ScienceDirect

The Journal of Finance and Data Science 8 (2022) 214–232
http://www.keaipublishing.com/en/journals/jfds/

A causal approach to test empirical capital structure regularities

Simone Cenci [a,*], Stephen Kealhofer [b]

[a] Leonardo Centre, Imperial College Business School, Exhibition Rd, London SW7 2BX, UK
[b] Blackstone Credit, San Francisco, 101 California St., CA, 94111, USA

Received 5 July 2022; revised 31 August 2022; accepted 9 September 2022
Available online 16 September 2022

Abstract

Capital structure theories are often formulated as causal narratives to explain which factors drive financing choices. These narratives are usually examined by estimating cross–sectional relations between leverage and its determinants. However, the limitations of causal inference from observational data are often overlooked. To address this issue, we use structural causal modeling to identify how classic determinants of leverage are causally linked to capital structure and how this causal structure influences the effect-estimation process. The results provide support for the causal role of variables that measure the potential for information asymmetry concerning firms' market values. Overall, our work provide a crucial step to connect capital structure theories with their empirical tests beyond simple correlations.

JEL classification: C1; C4; G3

Keywords: Causal inference; Structural causal models; Capital structure

1. Introduction

Understanding how firms finance their operations and which factors drive their choices is a fundamental issue in empirical finance. In the past 60 years, since the seminal work of Modigliani and Miller (MM)[1]; a plethora of different theories have been proposed to explain corporate financing behavior within a unified framework. Broadly speaking, we can group these theories into three categories. Static trade-off theories are based on considerations about balancing tax advantages against bankruptcy cost.[2] Considerations about costs associated with adverse selection due to an asymmetry in the information available to investors and managers, are at the heart of the pecking-order theories.[3] Finally, dynamic theories focus on the anticipation of the cost of rebalancing capital structures.[4,5]

In the past decades, researchers have examined the data to find support for, or reject, these different theories. Although there remain considerable differences as to the appropriate theory, there is increasingly a broadly accepted set of empirical findings related to firms' capital structures. These are nicely summarized in Frank and Goyal[6] under the

* Corresponding author.
E-mail address: s.cenci@imperial.ac.uk (S. Cenci).
Peer review under responsibility of KeAi.

https://doi.org/10.1016/j.jfds.2022.09.002

heading of stylized facts. One key set of the stylized facts relates the sign and magnitude of cross-sectional differences in leverage to a certain set of firm characteristics. Researchers have focused on the sign of these effects as different theories have been interpreted as leading to different signs. For example, one interpretation due to Frank and Goyal[6] is shown in Table 1.

We refer to this as the "contemporaneous correlation" model of capital structure. The explicit origins of the model lie with Rajan and Zingales[7] who first highlighted the general nature of the empirical relationships being observed. Of course, this is not a theoretical model of capital structure. In fact, these relationships have been exploited in a large number of empirical capital structure papers often with very different theoretical motivations. In fact, as they have been described, they are simply reliably observed statistical correlations.

In this paper, we are interested in two questions. (1) Are these effects causal or simple associations? (2) What is the magnitude of the causal effects and how do these effects relate to the associations found in the literature, i.e. do these variables actually matter? Answering these questions requires addressing two problems: *identification* of causal structures and *estimation* of causal effects.

As for any causal query, the ideal approach for answering our questions is a controlled experiment. However, because one cannot run controlled experiments with firms' capital structures, and natural experiment settings are rare, we need to use observational data. Causal inference from observational data has a long history that can be traced back to Wright[8]; Haavelmo[9]; Tinbergen et al[10]; Rubin.[11] Much of the causal inference literature in the social sciences and empirical finance is about estimation methods.[12] This includes methods such as regression discontinuity[13]; propensity score[14]; average and heterogeneous treatment effects[15,16]; and the potential outcome framework (PO).[17,18]

Methods to address the problem of identification of causal structures are significantly less common in the empirical finance research. The identification problem concerns testing ex-ante the causal structure that relates a variable to another (e.g., a particular firm characteristic to leverage), and, as we will show later in the text, it is a crucial step for correct estimations. A popular framework for addressing identification issues is the structural causal modeling (SCM) framework.[19,20]

In the SCM approach, we formulate a hypothesis concerning the data generating process and we represent it by a graph indicating directed causality, e.g. X causes Y: $X \rightarrow Y$. A causal graph is a Bayesian network (or probabilistic graphical model) that implies a set of conditional independencies among the node of the graph. The conditional independencies can be tested ex-ante on the data and can be used to derive the correct model specification to estimate unbiased causal effects. Broadly speaking, SCM provides a framework to answer our questions following four steps: (1) representing the causal story graphically by drawing a graph that encodes the assumptions about causal relationships; (2) testing these hypotheses by estimating the conditional independencies implied by the graph; (3) using the conditional independencies to select the necessary control variables and (4) estimating the effects with an appropriate estimator, e.g., regression, matching, propensity scores, causal forests. A detailed discussion of SCM for the reader not familiar with the topic can be found in the Appendix.

There are a number of differences between the causal empirical approach outlined above and standard statistical tests based on cross-sectional regressions. First, by estimating conditional independencies we can test our causal hypothesis ex-ante (i.e., before estimating the effects). That is, if our hypothesis concerning the causal relationship between two variables in the model is not supported, we can identify this misspecification explicitly and address it by changing the graph (i.e., model). Second, the theory of SCM clearly shows that to estimate unbiased causal effects, we should not control for every possible variable in the model, but only for those factors that confound the effect we are measuring. Importantly, SCM allows researchers to identify these factors.

The estimation of the relationship between firm size and market leverage provides an interesting empirical example that illustrates the problems associated with testing causal narratives with cross-sectional regression approaches that are

Table 1
Empirical capital structure regularities.

	Profitability	Market To Book	R&D	Tangibility	Selling Expense	Risk	Size
Pecking order	–	+	+	–	–	+	–
Trade off	–	–	–	+	–	–	+

The table shows the expected signs of the relationship between Leverage and a number of firms' characteristics according to the Pecking order (first row) and the Trade-Off (second row) theories of capital structure.

not based on validated causal models. In particular, the example highlights the importance of identification in causal inference from observational data. We discuss this example in the next section.

1.1. An empirical example: the case of size–leverage relationship

Consider the following controversy. Arguably, the trade-off theory implies a positive relationship between size and leverage: large firms are less risky, hence, for the same amount of debt have less default risk and therefore lower expected default costs. The pecking-order theory on the other hand has been interpreted to predict a negative relationship since larger firms are more established, face less adverse selection, and have a lower cost of equity issuance. What do the data say?

Numerous studies have found that size and leverage are positively correlated.[21,22] However Faulkender and Petersen[23] Hovakimian et al[24] and Hovakimian and Li[25] have found that, after controlling for whether or not a firm issues rated debt, the sign of the association may be reversed or vanishes. So, we have two theories (trade-off and pecking order) and two tests that provide evidence in support of both of them. The difference between the two tests comes down to whether we should control for rating choice when estimating the effect of size on leverage.

The decision as to whether or not a variable should be included in a regression model is typically based on considerations about the statistical performance of a particular specification.[6] However, SCM shows that the choice of whether to control or not for a given set of variables cannot be made by solely looking at associations in the data or other statistical measures. This is because controlling for variables that are spuriously correlated can increase model performance, but also introduce biases in the estimates of the effects. Without a causal hypothesis regarding the data-generating process, we cannot decide what we should control for to estimate unbiased effects. In this particular example, without a causal story that explains what drives firms' leverage, we cannot decide whether rating choice should be included or not in the regression and therefore we cannot conclusively establish the sign of the effect.

To illustrate that statistical measures alone can be misleading and lead to the wrong model selection consider the following toy model:

$$X = \mathcal{N}_x(0,1); \quad S = \mathcal{N}_s(0,1); \quad R = -0.5X + 1.5S + \mathcal{N}_r(0,1); \quad Y = S + \mathcal{N}_y(0,1) \tag{1}$$

where $\mathcal{N}(0,1)$ is the standard Gaussian. We only observe the data, not the model, and we want to measure the causal effect of X on Y. In the model the variable S plays the role of an unobserved covariate. From Eq. (1), it is obvious that this effect is zero. However we proceed here by first regressing Y on $X(y = \alpha x + \varepsilon)$ and look at the coefficient α of the regression, its standard error and the R^2 of the model. Then we add R to the regression and compare the statistics of the two models. A straightforward numerical simulation shows that, as expected, the coefficient from the first model is zero while the coefficient from the second model is not. However, the second model has a smaller standard error on the coefficient on X and a much larger R^2 (results are shown in Table 2). Therefore, it would typically be deemed to be the better model. Yet, we know by construction that this is not the case, i.e., the true effect of X on Y is zero.

In practice, it is standard to follow this two step procedure, i.e., starting with the regression of Y on X, and then considering adding other predictors, i.e., R. Indeed, as practitioners, we know that the t-statistic on X might be inflated due to correlation with R. If R does not meaningfully improve the overall fit of the regression, it is easy to discard R, taking its significant t-statistic to be spurious. But if it improves the regression fit, then normal practice would be to include it in the regression. This example shows that such an approach is wrong.

Table 2
Fallacies of standard statistical measures for model selection.

Covariates	Coefficient	Standard error	t-value	p-value	R^2	Residuals
X	−0.000	0.045	−0.009	0.502	0.001	Normal
X, R	0.229	0.037	6.099	0.000	0.346	Normal

The table shows the result of a regression from the simulation of the model in Eq. (1). The true effect of X on Y is zero, as correctly measured from the uncontrolled regression in the first raw of the table. The second regression yield a strongly biased estimate of the effect of X on Y but the quality of the fit as well as the confidence on the coefficient is significantly larger. This simple example clearly illustrates that statistical measures alone cannot help us determine the right model to estimate unbiased causal effects from observational data.

This simple model is designed to illustrate that, by fairly standard practice, we are going to draw the wrong inference regarding the effect of X on Y. In particular, statistical measures alone cannot provide enough evidence to distinguish between biased and unbiased effects (i.e., correct and misspecified models), and including more controls does not necessary lead to more accurate estimates of the coefficients. In fact, including the wrong controls not only artificially inflates the R^2, but it also potentially biases causal inferences.

Why does controlling for R in the previous example induces a bias? How can we identify the right structure of the model? SCM provides the tools to answer these questions. The causal graph associated with[a] Eq. (1) is $X \rightarrow R \leftarrow S \rightarrow Y$. In this example, the dependency between X and Y is blocked by the node R. After controlling for R we open up a channel (a spurious relation) between X and S and therefore X and Y. This spurious relation is called a *collider bias* because R is a collider node between X and S. In order to identify the right structure of the model and to avoid controlling for the wrong variables we need to formulate hypotheses (causal graphs) and test them ex-ante on the data. Going back to the empirical example, the only way we will be able to sort out the controversy regarding the causal effect of size on leverage is by determining where and how rating choice would enter a causal model of leverage through causal hypothesis testing. SCM provides a framework to formulate these tests.

1.2. Contribution and the organizational structure of the paper

In this paper we revisit the problem of estimating the effects of various firms characteristics on leverage ratios under a causal lens using structural causal modeling. SCM was first developed by Pearl[19]; Pearl et al[20] and has been used in the past to answer empirical questions in finance.[26–29] However, to the best of our knowledge our work is the first one to develop a method based on SCM to answer empirical capital structure questions. Specifically, our paper directly contributes to the empirical literature studying the relative importance of firms characteristics for the capital structure decisions of publicly traded firms[6,23–25,30,31]; Our approach is divided into three parts: (1) we specify a causal structure that incorporates the classic determinants of leverage; (2) we test for a causal role for those variables, and (3) we estimate the *total* effects of those variables found to have a causal role.

We find evidence of a causal role for size (+), profitability (−), enterprise risk (−), market-to-book (−), tangibility (+) and selling expense (−) for market leverage. The signs of these relationships are the same as those generally found in the literature Frank and Goyal,[32] but the magnitudes of the causal effects of these variables are mostly significantly larger.[23–25] We also propose a resolution to the size-leverage controversy discussed in section 1.1 and we find that the firm's decision to obtain a public debt rating should not be used as a control variable when estimating the causal effect of size on leverage. Overall, we believe that the structural causal model that we have developed for market leverage provides a useful point of reference for future empirical work on capital structure.

The paper is organized as follow: in section 2 we describe our dataset. In section 3 we replicate the approach of Hovakimian et al[24]; and Faulkender and Petersen[23] to illustrate the classic statistical analysis of capital structure regularities and the size-leverage controversy. We will use these results as benchmark for our causal approach. In section 4.1, we derive a causal model for leverage and we test the conditional independencies implied by the model. Then, in section 4.2, we use results from SCMs applied to our causal model, to derive a set of regression equations to determine unbiased estimations of the relationship between leverage and its determinants. Finally, in section 5 we discuss our results and compare our finding with the literature.

2. Data

In the following sections we will benchmark our results against those of Hovakimian and Li[25] and Hovakimian et al[24] (HKT) which are well-known and widely cited studies in empirical capital structure.[b] Therefore, we start by reproducing their dataset, with the main difference that we start and end at a later date, and we use quarterly rather than annual data.[c] We use data of North American companies from COMPUSTAT, and following HKT we remove financial firms (sic 6000–6999) and we only include firms with sales and book value of assets greater than $1 million. In order to

[a] A similar example to the one presented in this section can be found in Pearl et al[20]; pag 47.

[b] The two paper are qualitative similar but the latter has not been published. In the following we will denote both papers as HKT, unless there is a reason for differentiating the two.

[c] Results are robust if we use annual rather than quarterly data.

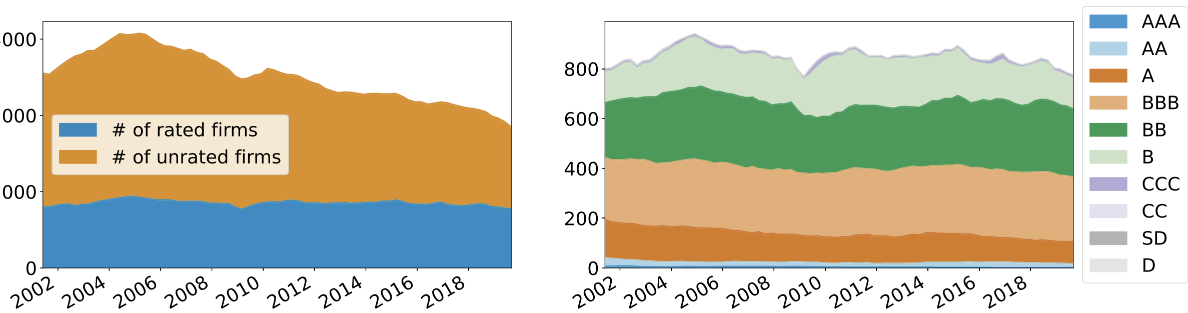

Fig. 1. **Overview of the population.** The left panel shows the count of firms per quarter in the whole population (orange) and in the rated population (blue). The right figure shows the composition of the rated population.

reduce the impact of outliers, we trim at the top one percent strictly positive ratio variables (e.g., Market To Book), and also the bottom one percent of ratio variables that can take on negative values (e.g., profitability). Fig. 1, left panel, shows a full count of the observations with non-missing values per quarter in the whole population (orange) and in the rated population (blue). The right panel shows the fraction of firms within each rating category. Table 3 provides summary statistics of the data in the observation period that goes from 2001 to 2019. For a direct comparison we present the summary statistics for rated and non-rated firms separately (this table can be compared with Table 2 in Hovakimian et al.[24] Ratings are the S&P implied senior unsecured long term debt ratings as reported by COMPUSTAT.

The main difference between our data and that of HKT is the time frame of the data. Our data runs from 2001 to the end of 2019; theirs from 1985 to 2008. Because the sample periods do not match, the quantitative results differ somewhat. Qualitatively, however, our data appears to closely resemble theirs, in that we can fairly reliably reproduce their results and the population characteristics appear similar (with the exception of Operating Risk). Following the definitions used in HKT, in Table 3 "probability rated" is the proportion of rated firms in the firm's sector. "S&P 500/400" and "NYSE indicator" are dummy variables set to 1 if a firm belongs to the respective S&P index or trades in the NYSE. "Market To Book" (MTB) is (book assets - book equity + market equity) over book assets. "Tangibility" is property, plant and equipment scaled by book assets. "R&D" and "Selling Expense" are research and development and selling, general and administrative expenses scaled by sales, respectively. Operating Income is defined as sales net of cost of goods sold and selling, general and administrative expenses. "Profitability" is Operating Income scaled by (one

Table 3
Summary statistics of the data.

Variable	Non Rated	Rated
S&P 500 indicator	0.03	0.39
S&P 400 indicator	0.06	0.18
NYSE indicator	0.18	0.68
Probability rated	0.14	0.20
Market to book	1.83	1.69
Tangibility	0.26	0.37
R&D	0.05	0.02
Selling Expense	0.31	0.17
Profitability	0.07	0.14
Size	4.50	7.68
Market Debt	0.12	0.26
Book Debt	0.26	0.37
Operating Risk	0.07	0.04
Market value of asset (log)	7.2	9.89
Volatility of asset	0.33	0.18
Observations	120,837	62,473

The tables shows the mean of the firm's characteristics within our population. The definition of the variables can be found in the main text.

Table 4
Parameter estimation from a statistical perspective.

	Rated population	Full population with rating control	Full population
Market to book	−0.054***	−0.02***	−0.026***
Tangibility	0.072***	0.112***	0.13***
R&D	−0.27***	−0.126***	−0.159***
R&D indicator	−0.033***	−0.025***	−0.03***
Selling expense	0.017	−0.047***	−0.05***
Profitability	−0.251***	−0.312***	−0.362***
Size	−0.015***	−0.006***	0.011***
Operating risk	−0.114	−0.01	0.062

The table shows the OLS coefficients of a linear regression model with market leverage as dependent variable. Standard errors are adjusted for heteroskedasticity and firm level clustering. *, **, *** denote significance at 10%,5% and 1%, respectively. The first column show the coefficients estimated over the rated population with the Heckman's correction. Except for Operating Risk, the values in the first column match the results found in Hovakimian, Kayhan, and Titman.[24] In particular, we have reproduced the negative relationship between Size and Leverage. The second column shows the coefficients from a regression over the full population with a rating control as in Faulkender and Petersen.[23] The third column shows the coefficients estimated using the full population without rating control. Notice that the coefficient on Size is positive and still statistically significant in this last regression.

quarter lagged book assets. “Size” is the log of sales adjusted for inflation. “Market leverage” and “book leverage” are (short-term debt + long-term debt) scaled by (book assets-book equity + market equity) and book assets, respectively. The term “Leverage” used alone should be interpreted as market leverage, and is the leverage variable of focus in this paper. “Operating Risk” is the standard deviation of Profitability measured over a rolling 4 year period.

The last two variables in Table 3 are proprietary measures for the market value of assets and its return volatility. These two quantities are derived as follows. Using the Merton model[33]; the market value and volatility of assets can be derived from the equity value, equity volatility and liabilities of the firm.[34,35] Specifically, by observing that the liabilities and equity are derivative securities of the underlying asset, one can relate the equity market value to the market value of the underlying firm assets by solving an inverse option problem, where liabilities are used to specify the boundary (“exercise”) conditions for the equity. A commonplace approximation in the literature is to measure the market asset value by subtracting book equity from book assets and add back the market value of equity.

Similarly, starting from the observed equity volatility we can solve an inverse option problem to estimate the volatility of the underlying market asset value. In practice, one solves a system of two equations, one for the value and one for the volatility of the asset, simultaneously. The averages of asset value and asset volatility, for the rated and non-rated populations respectively, are shown in the last two rows of Table 3.

3. Empirical capital structure regularities from a statistical perspective

Before presenting our approach we reproduce the results from Hovakimian et al[24] and from Faulkender and Petersen[23] using our updated data sample. The goal of this section is to provide a benchmark for our causal analysis Hovakimian et al[24] are primarily concerned with separating out the drivers of capital structure targets from factors causing deviations from targets. In doing this, they rely partly on credit ratings, however their study population includes both rated and unrated firms. In order to use the information in ratings, but control for differences between the rated and unrated populations, they condition on rating choice: whether the firm chooses to be rated or not. This should not be confused with a rating target: which rating the firm would like to have, if rated. Rather this is simply the choice of having any rating versus having no rating. Their regression procedure is a Heckman two stage process (Heckman correction). That is, they first run an independent Probit model for the probability of being rated using the full population. Then, they use the inverse Mills ratio estimated from this regression as an additional covariate in the second step regression over the rated population. For our purposes, they provide a Market Leverage regression which they report in their Table 4 under the rubric “Market Leverage Choice”.

In Hovakimian and Li[25]; they pursue a similar analysis, but in addition to ratings, they also use Moody's KMV EDFs, a commercially available estimate of default probabilities. The virtue of using the EDFs is that they can consider the unrated population without having to condition on rating choice.

In the following we run three analysis. First, we replicate the results of Hovakimian et al[24] Specifically, we regress Leverage on: Profitability, Operating Risk, Tangibility, Market To Book (MTB), R&D, an indicator variable that take a value of 0 if R&D is missing and 1 otherwise; Selling Expense, and Size. In this analysis we restrict the estimation only over the rated population, after controlling for self-selection bias with an independent Probit model.[d] Then, we reproduce the results of Faulkender and Petersen[23] which takes a somewhat similar approach, but controls for the rating choice using a dummy variable rather than via the Heckman correction. Finally, we run a regression over the full population without rating choice control. We run all the analyses over our full sample period using quarterly contemporaneous data.

The results of the regression are shown in Table 4. The first column shows the coefficients estimated over the rated population with the Heckman's correction (i.e., the HKT approach). Except for Operating Risk, the values in the first column match the results found in HKT. In particular, we have reproduced the negative relationship between Size and Leverage. The second column shows the coefficients from a regression over the full population with a dummy variable rating choice control as in Faulkender and Petersen.[23] Size and Leverage are still negatively correlated. The third column shows the coefficients estimated using the full population without any rating choice control. Here Size and Leverage are positively correlated.

In this section we have shown that with our dataset we can reproduce the classic empirical capital structure regularities observed in previous studies as well as the Size-Leverage controversy.

4. Empirical causal analysis

Here we revisit the classic empirical capital structure regularities from a causal perspective. This section is divided in three parts: first, we propose a causal model that describes the relationship between leverage and a firm's characteristics. Then we validate the model explicitly by testing the conditional independencies implied by the model. Second, we use the causal model and results from SCMs to estimate unbiased associations between leverage and its determinants. These are contrasted with the initial regression estimates. Finally, we revisit the Size-Leverage controversy discussed in section 1.1 through a causal lens. Understanding the results of this section requires some knowledge of structural causal modeling. In the Appendix we have provided a brief overview of the framework for the readers unfamiliar with the topic.

4.1. A structural causal model for the determinants of leverage

In this section we reformulate the model for Leverage using structural causal modeling. A structural causal model is a mathematical formalization of a causal narrative. Formally, a structural causal model is a set of structural equations with a well defined direction of causation. In SCM-ing, however, one makes minimal assumptions on the analytical form of the equations and, instead, focuses on analyzing the properties of the corresponding *causal graphs*. A causal graph is a graphical representation of the structural causal model. Each node in the graph is a random variable, and an arrow from one node X to another node Y indicates that Y is caused by X, i.e. Y can be written as a generic function of X, and possibly other variables. For each SCM, there exists one and only one causal graph and each graph entails a set of conditional independencies among variables. Because these conditional independencies can be validated ex-ante on the data (see Appendix, section D), a SCM can be, in principle, falsified.[e] This is in stark contrast with standard regression approaches where the model itself cannot be ex-ante falsified but only interpreted ex-post. Indeed, as we have shown in section 1.1 statistical measures alone, such as standard errors and goodness of fit of a regression model, cannot be used to select the *correct* causal models (see section B in the Appendix for further discussions). We see regression as an estimation method, not as a modeling tool.

In Fig. 2 we show the basic structure of our causal model for Leverage. In order to derive our model, we first identify what the key decisions for a firm are. Then we look at the effects of those decisions as well as at the factors driving them. Specifically, in our model a firm has two decisions to make over a given time interval (e.g., a fiscal quarter): (1) the desired amount of investment ("Investment decisions"), and (2) the desired external financing ("Financing"). These two variables are unobserved by outsiders and we have represented this unobservability property with a light gray color in the graph. Given the investment decision, the firm is faced with how to finance the investment. Given internally generated funds and the net debt issuance, the residual after investment is either paid out to or raised from equity. Thus

[d] Similarly to the referenced paper, we control for the probability of being rated as well as for the S&P 500 index indicator in the Probit model.
[e] As discussed in the Appendix, unobserved variables and identifiability problem limit the practical falsifiability of SCMs.

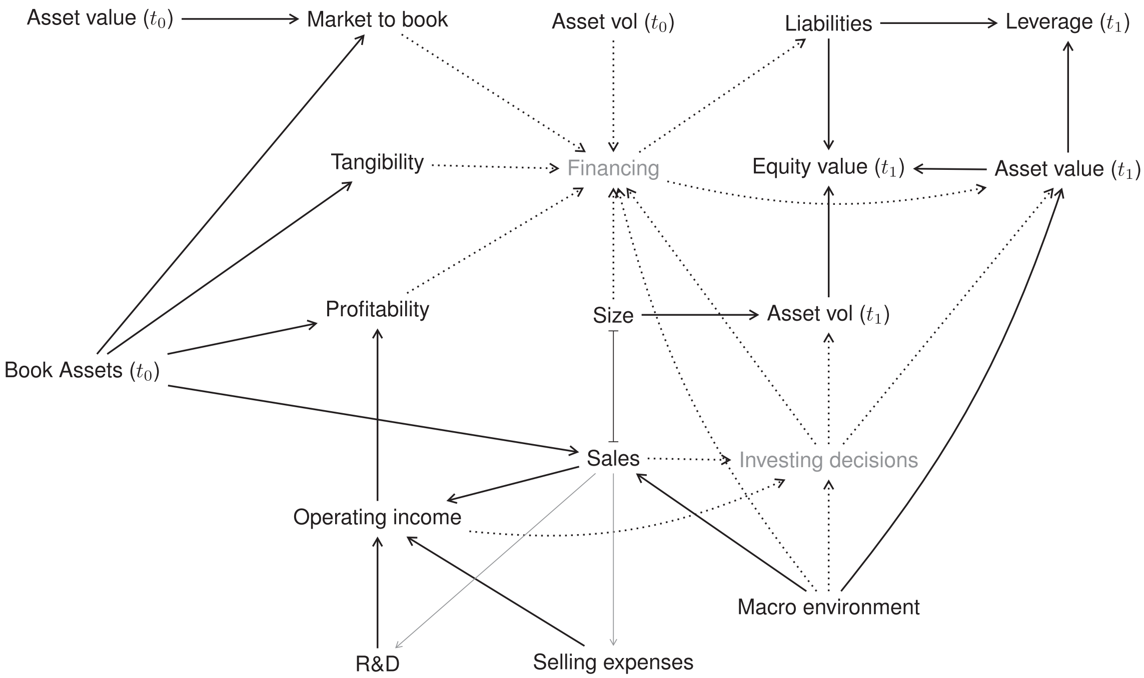

Fig. 2. **A structural causal model for the determinants of Leverage.** The figure shows our causal model for Leverage. Arrows indicate the direction of causation. The gray variables are unobservable investing and financing decisions, which are the two key decision of the firm. The dotted lines are the link to and from the unobservables. The gray thin arrows from sales to Selling Expense and R&D represent a link that is only there because of scaling reasons, without conveying additional information. Following standard notation, the idiosyncratic noise affecting each variable is omitted from the graph. The specific role of each node and link in the graph as well as the limitations in the empirical specification of this theoretical model are discussed in section 4.1.

the financing decision can be interpreted as either targeting equity payout (issuance) or net debt. The implication of capital structure targeting is that moving to the target in at least some circumstances benefits the firm, so the financing decision is understood to be a cause of asset value, which would also be reflected in the market value of the equity.

The investment decision cause changes in the asset value and asset volatility. The revised asset value, volatility, and observed liability level cause a new equity market value. The liability level and asset value cause a new Leverage level. Asset value and liabilities thus act as mediator variables for the effect of any other variable on Leverage. This structure follows the Merton model, treating equity as a call option on the firm's assets with boundary conditions determined by the obligation amounts of the liabilities.

Looking back at the factors driving the financing and investing decisions (the two key choices for the firm) we start from the assumption that a firm has existing assets and investment opportunities. These assets have an initial market value, volatility and book value (asset values (t_0), volatility (t_0), and book value (t_0) in the graph). The book assets generate sales which are also influenced by the macro-economic environment. The macro-environment could presumably encompasses a wide range of possible variables; we follow other capital structure research and merely use expected inflation and growth of GDP.[6] We assume all the common causes of the t_0 variables are fully captured by the realizations of their values and do not separately act on variables in the subsequent estimation period ("contemporaneous correlation" approach).

The temporal structure of the flow variables, such as sales, is deliberately under-specified. The context here is to treat these variables as causes of the financing and investing decision. Under this view, the decisions on investments or finance do not influence the values of these variables over the contemporaneous period. Conceptually, it is as if the liability changes are made discretely at the end of contemporaneous period. For variables like R&D and Selling Expense, this is potentially problematic, as these are flow variables presumably jointly determined with investments, another flow variable, and thus also impacting Operating Income. Our assumption here is that the contemporaneous period is short enough that these expenses are already largely determined. But this is an important assumption and limitation, particularly in a broader context of adjustment costs, lagged effects and distinguishing between target actions and causes of deviations from targets.

One could address this issue by distinguishing more explicitly between characteristics pre-determined at time t_0, such as past profitability, and treating contemporaneous profitability differently. That approach, however, does not conform to the “stylized facts” that we are addressing here, but would be useful for, say, the more complex analysis of target vs deviation from target. That is beyond the intent and scope of the present paper, and remains a topic for further research. In fact, the implicit assumption of the “market leverage choice” framework examined here is that the firm is continually striving to be at target so that the observed market leverage is simply a noisy version of the firm's market leverage target. We will revisit some of these points in the discussion section.

Overall, we believe that the framework presented in Fig. 2 is general enough to cover the structure of many of the cross-sectional regression models of Leverage in the literature (as surveyed in Frank and Goyal[6]; for example, and represented under their “stylized fact 8”). However, it is also important to note that the actual variables used in the specification of this model (or any model) may not correspond to the conceptual quantities they are intended to represent. For example, there are multiple ways of specifying firm size; presumably some might produce different results and thus undermine the generality of the economic interpretations of the findings. This issue cannot be eliminated but our objective can be defined more narrowly, namely to examine the effect of variables as conventionally specified and interpreted in the existing literature. To this end we follow the standard variable definitions used in most of the literature.[6,24,25] Another important problem is under-reporting, such as for example in R&D expenses. As we will see in the next section under-reporting leads us to drop this variable from the empirical specification of the model. In section 5.1 we discuss further limitations of the model and the applicability of SCM to capital structure research in general.

Now that we have presented our framework using a structural causal model, we endeavor to validate the hypotheses by testing the conditional independencies implied by the graph.

4.1.1. Conditional independence tests

To test if there exists a causal path from a firm's characteristics into Leverage we need to test the conditional independencies implied by the model in Fig. 2, or alternatively, we need to test the counterfactual hypothesis of a graph without the link that we have assumed exists (the counterfactual approach will be explained in more detail below in the context of an example.). The conditional independencies implied by the graph can be recovered by applying definition A.1 (d-separation) to the model in Fig. 2. The independence tests are our tools to validate the model, and a central part of our methodology. Therefore, it is important to understand how they work. The goal of a conditional independence test is to assess whether a variable (X) is independent of another variable (Y) given a set of variables (Z), i.e., $X \perp\!\!\!\perp Y | Z$. If two variables X and Y are unconditionally dependent there could be a (direct or indirect) causal link between the two. However, if the conditioning set is not in the path from X to Y and it removes this dependency, then their association is spurious and there is not causal link from X to Y (or Y to X). Otherwise, provided that we observed all the factors in Z, the link exists.

The general procedure to measure conditional independencies consists in (1) regressing (either parametrically or non-parametrically) X on Z and Y on Z, then (2) testing the independence of the residuals with a t-test or a kernel based test.[36] Kernel conditional independence tests (KCIT) offer the most flexible solution to the problem of measuring conditional independencies because they rely on minimal assumptions on the data generating process. However, they can be computationally expensive. Here we use the implementation of KCIT developed in Zhang et al[37] and Strobl et al.[38] Broadly speaking, the algorithm proposed in Strobl et al[38] (the randomized conditional independence test, RCIT) is an hypothesis test on the norm of the partial cross-covariance matrix of (X,Y) given Z (see Appendix, section D for further details). Therefore, the output of the RCIT is a p-value. To analyze the result of the independence tests we compare the distribution of p-values of unconditional and conditional tests as follow. If two random variables are (conditionally or unconditionally) independent, then the p-value of a number of independent independence tests is uniformly distributed between zero and one. If the random variables are dependent instead, the distribution of p-values will not be uniform but rather concentrated around zero (see Appendix, section D for an example using synthetic data). Therefore, by running the independence tests over multiple independent cross sections and comparing the conditional and unconditional distributions we can determine whether the graph is compatible or not with the data.

In the following we run all the conditional independence tests cross-sectionally in (percentage) quarterly changes. That is, we test if changes in a variable X are conditionally independent on changes of a variable Y, given changes in the variables within a set Z (i.e., $X \perp\!\!\!\perp Y|Z$). To increase the power of the test we increase the sample size by splitting each cross-section into three subsamples with approximately 1000 (independent) observations each. This splitting effectively increases the size of the overall sample by a factor of three.

The two main tests for uniformity of empirical distributions are the Kolmogorov–Smirnov test and the χ^2-test. Because the former requires large samples, but we run the conditional independence tests on relatively small samples, here we use the latter. Specifically, we report the p-value of the test of the unconditional and conditional distribution. If the p-value of the test is smaller than 0.1 (non-uniform) in the unconditional setting and greater than 0.1 (uniform) in the conditional setting, then conditioning remove the dependencies between the two random variables. In this case, the association between the two is considered to be spurious. Otherwise, the causal association of the SCM is considered to be consistent with the data. In the following we test the structure of the model using the conditional independence tests by looking at each determinant of Leverage independently.

4.1.1.1. Profitability. Profitability is defined here as Operating Income over total book assets as of the start of the interval (t_0). We can use Profitability as an example to illustrate the application of d-separation (see section A, definition A.1). Broadly speaking, d-separation is the concept in graph theory that allow us to determine the implied conditional independencies of the model. In the model in Fig. 2, Profitability and Leverage are d-separated only by asset value and liabilities. Therefore the only conditional independence test we are able to run is Profitability $\perp\!\!\!\perp$ Leverage | asset value, liabilities. However, this is a trivial conditional independence test as Leverage is a deterministic function of the

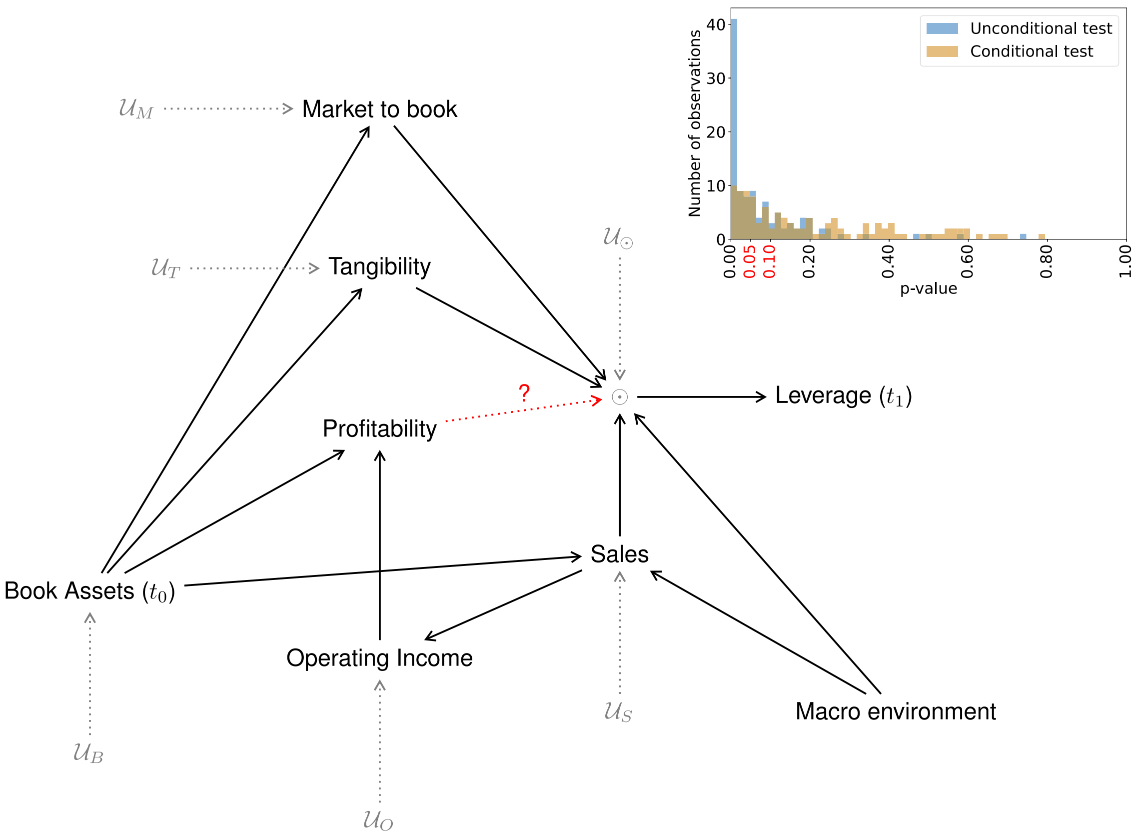

Fig. 3. **Relationship between Profitability and Leverage.** The figure shows the graph associated with the conditional independence test that we run to validate the role of Profitability in the model. The gray nodes represent idiosyncratic drivers of the variables shown in the graph. The circle represents all the possible mediators of the effect on Leverage. The goal of the test is to assess whether the red dotted link exists. The inset shows the distribution of the conditional and unconditional test and it illustrates that after controlling for all the non-causal channels, Profitability and Leverage are still conditionally dependent, although, as expected, part of the dependence is removed.

conditioning set (see section D for a discussion of the role of determinism in structural causal models), i.e., Leverage is fully determined by asset value and liabilities.

In order to test our hypothesis concerning the role of Profitability we take a counterfactual approach. That is, we show that, if the link Profitability → ⊙ → Leverage does not exists, the conditional independence test implied by this counterfactual graph are not validated by the data (here ⊙ is the set of mediating factors of the effect of Profitability on Leverage). Specifically, if we remove the link Profitability → Financing, then the following conditional independencies must be true: (1) Leverage ⫫ Profitability | Book asset, Operating Income (2) Leverage ⫫ Profitability | Operating Income, MTB, Macro environment, Sales, Tangibility. Condition (1) is again trivial because Profitability is a deterministic function of the conditioning set. Test (2) is not trivial, and it says that if Profitability has no (indirect) link to Leverage, then the (spurious) correlation should be removed after controlling for all the other non-causal channels. Fig. 3 shows the graph representing the conditional independence test we are after. Because (2) is the only non-trivial test it will be the only one we run. The results of the test are shown in Table 5. The table show that after conditioning, Leverage and Profitability are still not statistically independent, therefore the data support the hypothesis that the link Profitability → ⊙ → Leverage exists.

4.1.1.2. Market To Book. Following standard definitions, MTB is the ratio of the asset value over book assets.[24,30,39] There is extensive evidence in the literature that documents the existence of the link between MTB and Leverage Baker and Wurgler (2002).[30,40,41]; The trade-off and the pecking order theories are usually interpreted to predict opposite signs for this relation (negative and positively, respectively).

By construction MTB is driven by both the asset value (t_0) and book assets (t_0). Therefore, following model 2 and definition A.1 if the link from MTB into Leverage does not exists then both the following must be true: (1) MTB ⫫ Leverage | Book assets, and (2) MTB ⫫ Leverage | Profitability, Macro environment, Sales, Tangibility. Table 5 rows two and three, shows that after conditioning on both sets, MTB and Leverage are still dependent.

4.1.1.3. Nature of assets: R&D, tangibility and selling expense. Selling Expense (selling, general and administrative expenses, XSGA, over sales) has a path into Leverage through Operating Income. This link is an accounting identity. Both the pecking order and the trade-off theory are interpreted to predict a negative relationship between Leverage and Selling Expense. Table 5 shows that the one path through Operating Income is not the only path of Selling Expense into

Table 5
Hypothesis testing on the causal graphs.

Hypothesis	χ^2
Profitability ⫫Leverage| Sales,Op. Inc.,Market to book, Tangibility,Macro environment	0.0 → 0.04
Market to book ⫫ Leverage| Book assetst_0	0.0 → 0.0
Market to book ⫫Leverage| Sales,Op. Inc.,Profitability, Tangibility,Macro environment	0.0 → 0.0
Selling Expense ⫫ Leverage| Sales	0.0 → 0.0
Selling Expense ⫫ Leverage| Sales,Op. Inc.	0.0 → 0.0
R&D ⫫ Leverage| Sales	0.0 → 0.26
R&D ⫫ Leverage| Sales,Op. Inc.	0.0 → 0.26
Tangibility ⫫ Leverage| Book assetst_0	0.0 → 0.0
Tangibility ⫫ Leverage| Sales,Op. Inc.,Profitability, Market to book, Macro environment	0.0 → 0.0
OpRisk⫫ Leverage| Op. Inc.	0.41 → 0.61
Sales ⫫ Leverage| Op. Inc.,Profitability, Market to book, Tangibility,Macro environment	0.0 → 0.0
Sales ⫫ Leverage| Op. Inc.,Book assetst_0,Macro environment	0.0 → 0.0
Risk ⫫ Leverage| Asset valuet_0,Book assetst_0	0.0 → 0.0

The table shows the result of the conditional independence tests. The distribution of the p-values are shown in figure S5. The first column shows the hypothesis. For each hypothesis we show the results of the unconditional and → conditional test. The second column is the p-value of the χ^2-test after binning. If p-value > 0.1 the distribution of p-values of the independence test is uniform, therefore the variables are (conditionally or unconditionally) independent. For example, the first row show that Profitability and Leverage are unconditionally dependent, and still dependent after controlling for the conditioning set. On the other hand, the table show that Selling Expense is conditional independent on Leverage after controlling for Size and Operating Income. Overall, the table validate the model shown in Fig. 2.

Leverage as also after conditioning the two are still conditionally dependent. This supports the existence of the causal relationship.

Research and development is the ratio of the book value of research and development expenses (XRD) and sales. In the literature, R&D expenses are often included as drivers of Leverage because they are associated with both proprietary production and growth, and thus identify firms with equity values prone to adverse selection (an important hypothesized driver of the Leverage target). R&D expenses are subtracted from revenue in the calculation of Operating Income. Therefore, in the causal model R&D has a path into Leverage through its effect on Operating Income.

It is important to stress that this link is an accounting identity. R&D is a long-term investment that presumably should not result in any benefit to the firm in the period of investment. Therefore its true causal effect cannot be measured contemporaneously. Table 5 shows that our data do not support the hypothesis proposed in the SCM graph. We believe that this negative result from the conditional independence test, given the nature of the accounting identity being examined, is most likely due to a data issue. Many firms do not report research and development expenses (~50% do not in our sample). For this reason HTK (in line with common practice) also used an indicator dummy variables for unreported R&D in their regression specification. However, because we are unable to conclusively show that the source of the negative result is a data issue, in the following analyses we exclude R&D as a causal variable.

Tangibility is property plant and equipment (PPE) over book assets. Here we hypothesis that the effect of PPE on Leverage is mediated by the financing choices. Therefore, the conditional independencies to test are similar to those of MTB and are shown in Table 5. The results support the hypothesis that Tangibility drives Leverage.

4.1.1.4. Risk. Similarly to the case of Market To Book, the pecking order theory and the trade-off theory predict different signs for the effect of risk on Leverage. Yet, in both theories risk is an important driver of Leverage Frank and Goyal (2009). Here we consider two different definitions of risk. In most of the literature, risk is measured as the volatility of accounting income. For example, in HKT Operating Risk is the volatility of the Profitability measured over a 4–5 year period. The problem with such measures is that they tend to be quite noisy, due to the limitations of accounting data. For instance, using annual data, one would only have 4 to 5 observations to determine a standard deviation. Using quarterly data, there are marked seasonal effects for many firms. Using Operating Risk as described above and conditioning on Operating Income, we have found that the data do not support the hypothesis that this variable is a causal driver of Leverage (see Table 5)). In fact, the dependence between Leverage and Operating Risk is also dubious unconditionally.

As an alternative to Operating Risk, we define risk as asset volatility, the standard deviation of the return to the asset value. In our causal model, the contemporaneous volatility of the asset value is a driver of equity value (per the Merton model), but not asset value and thus not Leverage, i.e., there is an underlying process that drives the value and the volatility of the assets but only the former is a determinant of Leverage. However, the asset volatility at t_0 is a driver of the financing decisions and therefore drives Leverage through an effect on liabilities. Because we have assumed that the only t_0-relevant variables are the market and book asset values and volatility, these are the only variables on which we need to condition (i.e., if Leverage and risk are still dependent after blocking these spurious paths then the link exists). The tests support the causal link.

4.1.1.5. Size. We measure Size as the log of sales adjusted for inflation as per HKT. . As discussed in section 1, the trade-off theory and the pecking order are interpreted as predicting opposite sign for the relationship between Size and Leverage (positive and negative, respectively). Macroeconomic conditions are also important determinants of Leverage Frank and Goyal (2009) through their effect on sales, financing and investing decisions. Here we use expected inflation, growth in GDP and interest rates as macroeconomic factors Frank and Goyal (2009).

The hypotheses shown in the graph are implicitly based on several assumptions. Size is proportional to the number of projects a firm can take Vuolteenaho[42]; hence it directly drives the volatility of the assets, i.e. larger firms can be seen as well-diversified portfolios of investment projects. Structurally, this also implies less potential for asymmetric information as diversification reduces the proportion of idiosyncratic risk. Moreover, Size also drives the financing choice as variation in revenue presumably influences the investment decision. Because the Size variable is essentially the same as the sales variable, it is not possible to condition on sales in evaluating the conditional dependence of Size and Leverage. That means that we need to consider all the paths by which sales has an influence. Similarly, to the previous analysis we will therefore validate the graph counterfactually: if sales were not to drive investing decisions and financing choices, then all the effects on Leverage would be those mediated by Operating Income and those spuriously produced by book assets at t_0. Therefore, Leverage and sales would be independent after conditioning on the macro environment, Profitability, Tangibility and MTB: sales $\perp\!\!\!\perp$ Leverage | MTB, macro environment, Profitability, Tangibility. The result of these tests are shown in Table 5 and indicate that sales does indeed drive Leverage not only through Operating Income (and Profitability) but also through financing and investing choices.

4.2. Parameters estimation from the causal model

In section 3 we reproduced the results of HKT. Then, in sections 4.1 and 4.1.1 we have derived and empirically validated a structural causal model for the relationship between leverage and firms' characteristics. Here we use the model to (1) revisit the sign and significance of these relationships and (2) address explicitly the size-leverage issue discussed in the introduction.

A simple application of the backdoor criterion (see Appendix section A definition A.2) to the graph derived in the previous section shows that, by including all variables within a single regression model, we induce a bias in the estimated coefficients. As an example, consider the relationship between Selling Expense and Leverage. Selling Expense have a causal path into Leverage through Operating Income. Operating Income drives Profitability which is a descendant of a mediator variable for the effect of Selling Expense on Leverage. Hence if we include both Profitability and Selling Expense in a model for Leverage, the coefficient on Selling Expense will not measure the total effect anymore, but rather a covariate specific effect (see section B.2 for a detailed discussion of how this control can change the value of the total effect).

Using results from SCM we can now turn to the general results. Table 6, shows the parameters estimated for each variable separately including all the necessary and sufficient controls as derived from our causal model using results from SCM (see Appendix). The control sets are shown in Table 6 second columns. Specifically, to estimate the effect of MTB we control for Operating Income (log), Profitability, Tangibility and Size. To estimate the effect of Tangibility we control for Operating Income, Profitability, MTB, and Size. To estimate the effect of Selling Expense we only control for sales (log). We do not estimate the effect of R&D because we could not validate its role in the model. To estimate the effect of Profitability we control for Operating Income, MTB, Tangibility and Size. To estimate the effect of Size we control for book asset (log) and the macro environment (expected inflation, GDP growth, and the level of the 10 years treasury). Finally, to estimate the effect of risk we control for the book asset and market asset value at t_0. Notice that, as discussed in section 4.1, following our model, the market and book assets value are the only other t_0-relevant variables. The goal of these regressions is to estimate unbiased total linear effects.

Overall, compared with Table 4 we find that the effects have the same sign as the stylized facts of Frank and Goyal.[6] A key difference with the full sample results of HKT and Faulkender and Petersen[23] is in the magnitude of the estimated coefficients. We also note that HKT find a very significant positive relationship between Operating Risk and Leverage, while we find a strong negative association between our measure of risk (asset volatility) and Leverage. Using a conceptually similar measure of risk to the one we use,[23] also find a significant and negative relationship between risk and leverage.

In contrast to regression, which is simply identifying the statistical associations that exist in the data, the causal estimates are constructed to estimate the total causal effect on leverage of varying a given variable. To the extent that the causal variables are themselves related, the causal estimates will generally differ from regression estimates that include

Table 6
Parameter estimation from the causal model.

	Causal effect	Control set
Market to book	−0.06***	Op. Inc.,Profitability, Tangibility, Size, Macro
Tangibility	0.176***	Op. Inc., Profitability, Market to book, Size, Macro
Selling expense	−0.34***	Size
Profitability	−0.362***	Op. Inc., Market to book, Tangibility, Size, Macro
Size	0.046*	Book assett_0, Macro
Risk	−0.926***	Book assett_0, Asset valuet_0

The table shows the OLS coefficients of a linear regression model with market leverage as dependent variable. Standard errors are adjusted for heteroskedasticity and firm level clustering. *, **, *** denote significance at 10%,5% and 1%, respectively.

Table 7
Economic significance of the coefficients.

	Statistical model (with rating control)	Statistical model (without rating control)	Causal model
Market to book	−0.030	−0.039	−0.089
Tangibility	0.031	0.036	0.048
Selling expense	−0.013	−0.014	−0.093
Profitability	−0.049	−0.057	−0.057
Size	−0.012	0.022	0.093
Operating risk	−0.000	0.002	
Risk			−0.130

The table shows the difference in the economic significance of the coefficients in the statistical and causal model. Specifically, we show the effect on Leverage of one standard deviation move of the independent variables. The economic significance is estimated for the full sample with a rating control variable (first column), the full sample without rating control (second column), and the full sample with the causal model (third column). Because we define Risk differently from the literature we present the results in a different row. Overall the table shows that the causal model suggests a higher economic significance of the determinants of Leverage.

all of the variables simultaneously. The general finding here is that the causal effects are significantly larger in magnitude than those estimated via conventional regression methods.

To appreciate the significance of these differences, we show in Table 7 the effect on Leverage of a one standard deviation move, first using the HKT full sample estimates with and without rating control (as in the second and third column in Table 4), and second the SCM estimates (as in Table 6). Table 7 highlights the economic significance of the coefficients estimated from the causal model (third column), especially when compared to the coefficients obtained after controlling for rating choice (first column).

Finally, using our causal model and the estimation method discussed in this section we can revisit the size–leverage relationship controversy discussed in the introduction. Specifically, we are interested in the question: Should we control for whether or not a firm issues rated debt in order to estimate the unbiased association between Size and Leverage? HKT control for rating because they are interested in exploiting the capital structure target information in ratings, based on the rated population. Issuing rated debt is a voluntary decision and therefore there is a selection bias in their sample, which they correct for with the Heckman procedure. On the other hand,[23] control for rating over the full population without sample bias corrections. In their specification, rating choice is a binary variable. Is this a correct control? To answer this question we need to understand how rating choice enters into the model in Fig. 2.

Fig. 4 panel (a) illustrates our hypothesis. Rating choice is effectively a financing choice, and it is driven by the size of the firm, as there are well known constraints on debt issuance amounts that keep small firms from accessing public debt markets. Rating choice also depends on the total amount of debt. Indeed, it has already been shown that firms with long-term debt and publicly traded equity generally are more likely to issue rated debt.[23] Other factors driving rating choice are assumed to include the overall financing decision. For this analysis, we do not need to be explicit about these dependencies and therefore we include them in the unobserved variable. The causal graph provides us with an explanation for the size–leverage relationship controversy. Under this graph, controlling for rating choice induces a bias

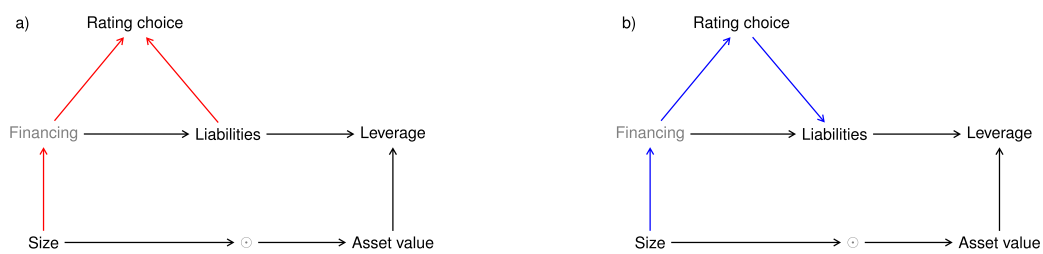

Fig. 4. **Role of rating choice in the causal model.** The figure shows two hypothesis for the role of rating choice in the model. In a) we assume that rating choice is driven by both Size and the total liabilities. In b) we assume that rating choice is driven by Size and drives the total liabilities. In both scenarios, we should not control for rating choice to estimate the total effect of Size on Leverage.

in the estimated effect because it is the central node of a collider: size → rating choice ← debt. Recall that controlling for a middle node of a collider induces a spurious correlation between the other two nodes. If our model is correct then the controversy in the size–leverage relationship, i.e., the sign difference in the specification with and without rating choice control, is due to a spurious correlation induced by a collider bias.

An alternative hypothesis for the role of the rating choice in the graph is that the choice to issue rated or unrated debt drives the total amount of liabilities (see Fig. 4 panel (b)). If this was the case, then the rating choice would not be the central node of a collider. However, it would be the descendant of a mediator variable and as shown in section B.2 controlling for it would lead to measures of covariate specific, not total, effects. Covariate specific effects are different to total effects in that part of the effect is blocked by controlling for one of the channels in the graph from Size to Leverage (see Fig. 4 panel (b)). Under this hypothesis therefore the negative sign induced by the rating control is one of the many possible covariate specific effects of Size on Leverage, not the total effect.

Another possible relationship is reverse causation between rating choice and Size. In this account, a public debt rating reduces financing constraints and/or lowers capital costs, enabling the firm to grow larger. Although a plausible relationship, it does not accord with the contemporaneous correlation framework. Just as with R&D expense, it is difficult to argue that the effects would show up coincident with the rating choice action.

Notice that, differently from the other variables, rating choice is a binary variable. Unfortunately, measuring conditional independencies for mixed discrete and continuous variables is notoriously challenging.[43] Therefore, we are not able to validate which one of the two hypothesis in Fig. 4 is supported by the data using conditional independence tests. However, we do not need to decide between these two hypotheses in order to reject that we should control for rating choice when estimating the total effect of Size on Leverage. In particular, the financing decision does not drive Size (or any of its ancestors). Therefore, rating choice is not a plausible confounder of the effect of Size on Leverage, regardless of the exact position of rating choice in the causal model. Moreover, selection bias is not relevant here because we observe Market Leverage for all firms in our sample. In sum, following our model, to estimate the total effect of Size on Leverage we should not control for rating choice.

5. Discussion

Understanding the causes of differences in firms' capital structures is important in resolving a number of empirical issues from explaining cross-sectional variations in credit spread to understanding financing constraints. The two main theories of capital structure are the trade-off theory and the pecking order. Quoting Frank and Goyal[6]; these theories "provide points of view", providing "guidance for the development of models and tests. But neither is tied to a specific model formulation." As a consequence, the implications of the models have often been interpreted differently. For example, there is not agreement on how certain empirical variables map to the characteristics described in the models. Does profitability measure growth prospects or does it measure sufficiency of cash flow? Or intangibility? Or all three? One synthesis that Frank and Goyal[6] have drawn from the literature is shown in Table 1.

On the other hand, there is more agreement around certain of the empirical regularities related to capital structure. Beginning with Rajan and Zingales[7]; and well summarized, again, by Frank and Goyal[6] in their stylized fact 8: "There is a core set of reliable factors that are correlated with cross-sectional differences in leverage". Specifically, leverage is positively related to: median industry leverage, collateral (tangibility), log of sales (size), and expected inflation; leverage is negatively related to market-to-book and profits. This set of variables is frequently augmented by additional generally accepted explanators of leverage: research and development expenditure (−); selling, general and administrative expenses (−), and measures of firm risk (−).[23,25]

Since Rajan and Zingales[7] pointed out these variables, they have appeared as determinants of leverage in a variety of papers, papers often motivated by quite different explanations of capital structure (see for example Baker and Wurgler[4]; Frank and Goyal[6]; Grahamand Leary[22]; Fama and French.[30] Our objectives here are to see if these variables can be validated as causal variables in the spirit of Pearl et al[20]; and to measure their causal effect on market leverage.

If we contrast our results with HKT, there are three primary differences. One of the primary differences relates to the measurement of risk. We find that the risk measure that they employ does not have a causal relationship to leverage in our data. We also cannot reproduce the statistically significant coefficient that they find in their regression studies. If we use a measure of market asset volatility rather than cash flow volatility, we find support for a causal relationship and a strong negative association with market leverage, in fact one of the strongest economic effects in the data. This is consistent with the findings of Faulkender and Petersen[23] who use a measure of risk similar to the one we use.

A second difference relates to the sign of the Size effect. As per Faulkender and Petersen[23]; they find an ambiguous sign effect. Notably, when they impose rating choice as a control variable, they find that the sign of the Size effect switches from positive to negative. Similarly Hovakimian et al[24] have found that the rating control also induces a negative sign, and Hovakimian and Li[25] have found that after controlling for rating choice, the effect of size goes to zero. We cannot resolve this issue empirically, due to technical difficulties in testing for conditional independence with a mix of discrete and continuous variables, but we have shown in the causal framework that it is difficult to support rating choice as a necessary control variable. In fact, the most likely explanation of the observed effect is that rating choice is inducing a "collider" bias in the estimation of the Size effect.[f]

The other notable difference with the full-sample results is in the magnitude and economic significance of the variables. The SCM approach suggests a more significant economic role for some of the commonly employed capital structure variables than is evident from the cross–sectional correlation approaches. In fact, overall, for the variables that show evidence of a causal relationship (Size, MTB, Tangibility, Profitability, Selling Expense, and asset volatility), we generally find significantly larger effects than those obtained via the standard regression methods. For instance, we find that a one standard deviation move in Size produces a 0.093 move in market leverage (versus 0.022). A one standard deviation move in MTB produces a (negative) 0.089 move in market leverage (versus 0.039). A one standard deviation increase in asset volatility reduces leverage by 0.13.[g]

The differences in the economic significance is even greater when compared to the full-sample model with the rating control. When comparing the economic significance of each variable with those found in the literature, however, it is important to notice that here we focused specifically on total effects. The coefficients in the regression models encountered in the literature instead are a mix of covariate specific and (possibly biased) total effects. Regression based approaches do not explicitely distinguish between the nature of these effects, because making this distinction requires an understanding of the causal structure of the data generating process, and the specific role of each covariate. On the other hand, SCM provide a framework to distinguish total from covariate specific effects. Clearly distinguishing between the nature of the effects measured in different regression specifications is crucial for comparing the relative importance of different factors, and to interpret the coefficients in relation to the original research question or hypothesis.

Comparing our findings (Table 5), with Table 1, one interpretation is that we have found strong support for the trade-off theory. However, in our view, the findings are supportive of dynamic theories of capital structure. Those theories focus on the anticipation of recapitalization costs, for instance, the cost of raising equity in difficult times. There is considerable evidence from equity issuance on the role of asymmetric information and adverse selection in determining those costs. Firm characteristics that give rise to greater potential for asymmetric information, namely higher levels of idiosyncratic risk, are associated with unrealized growth opportunities, uniqueness of product line, as well as overall levels of firm risk. The variables highlighted here have been identified with those characteristics. Market-To-Book predicts future growth. Profitability and Selling Expense are related to uniqueness of product line. Idiosyncratic risk increases with total risk. Tangibility and Size are both negatively associated with idiosyncratic risk.

5.1. Limitations of SCM in capital structure research

We now briefly discuss a number of important limitations for the use of SCM in empirical capital structure research.

One principal limitation of SCM in this application relates to simultaneous determination Imbens (2019). The framework does not apply when the underlying graphic is cyclic, i.e. $X \rightarrow Y \rightarrow X$. Some of the relationships in the model we propose ignore this possibility. For example, it seems quite reasonable to maintain that sales over the contemporaneous interval are not caused by the investment decision over the same interval. But once we begin looking at variables such as R&D expense, the direction of causation over the contemporaneous interval could be reversed (from investment decision to R&D expense). In this case, it may be impossible to disentangle the "causal" effect of R&D expense on Leverage, or even the effect on liabilities alone within a SCM framework.

[f] We would like to note that the potential bias that we have identified in their analysis arises due to their focus on using information in ratings to identify capital structure targeting behavior and the econometric problems which are presented by unrated firms. This is a more ambitious analysis than what we have undertaken here. In the large, we find their results credible and significant, which is why we have used them as a baseline for this analysis.

[g] As a reference point the 25th and 75th percentile of leverage is 0.1 and 0.45, respectively[44].

The estimates vulnerable to this criticism are those associated with current cashflow: Profitability, R&D and Selling Expense. The problem arises from a reversal in the direction of causality. If, for instance, the firm has Leverage above target and acts to reduce or repay debt by cutting R&D or Selling Expense, this will induce a positive association between these variables and Leverage. Reducing these variables would also raise Profitability, inducing a negative association with Leverage. The causal estimates for these variables, assuming no reverse causation, are both negative and large, suggesting at least in the case of Selling Expense that reverse causation does not appear to be a problem. Since any reverse causal effect on Profitability would have to work via Selling Expense, this is also indirect evidence as well against reverse causation in that channel.

Another limitation is that the SCM framework does not eliminate that the direction of causality could be from the enterprise value to the financing decision. Under this scenario, for instance, a firm whose market enterprise value goes up has lower leverage, but also the increase in enterprise value could cause increased investment via higher selling expense or could be anticipating an increase in profitability. However, we would argue that, under the assumption that the firm is continuously targeting its market leverage, the key issue is why the firm would allow its leverage to go down. In other words, if our focus is on market leverage, the influence of enterprise value on other causal variables is not relevant, as long as the firm is in the position to offset the shift in enterprise value to maintain target leverage. This argument breaks down if there are reasons, such as adjustment costs, for the firm to delay responding to changes away from target leverage. But this is a general critique of the "contemporaneous correlation" model of capital structure.

Finally, an important limitation on the economic interpretation of causal effects is related to the interpretation placed upon a particular empirical variable. Let us consider the case of Profitability. The conventional interpretation is that this measures how profitable the firm is, and this in turn is an indication of some degree of market power due to uniqueness of its product(s). However, it is not difficult to put other interpretations on this variable, generally related to mis-measurement of book assets. Demonstrating a causal linkage between Profitability and Leverage is therefore amenable to other economic interpretations. We do not resolve such issues. In this analysis we have simply used conventional variable definitions, so that our results can be interpreted and compared with existing works.

While considering these limitations, it is important to bear in mind that they also apply to virtually all empirical capital structure research. Structural causal modeling does not resolve these issues. SCM provides an explicit approach to the causal relationships in an empirical model. As we have shown here, this helps in avoiding spurious associations and in determining the actual magnitudes of the posited causal effects.

6. Conclusion

Capital structure research has converged in recent decades on a set of empirical regularities that are often cited and regularly included in model specifications. A subset of these could be characterized as the contemporaneous correlation model of (market) leverage. These are a group of primarily firm accounting characteristics that have been connected with fundamental characteristics such as expected growth rate or risk. In this paper we attempted to answer two questions: Do these characteristics cause capital structure or are they merely associations? If we can establish a causal link, what is the direction and economic significance of the effect?

We used a methodology based on a well-established causal inference framework: structural causal modeling (SCM). We derived and explicitly validated a causal model for market leverage and used results from SCM to estimate the relationship between firms' characteristics and market leverage ratios under a causal lens. Overall we have shown that Profitability, Market-To-Book, Selling Expense and market asset volatility are economically and statistically significantly and negatively signed causes of market leverage; Size and Tangibility on the other hand are economically and statistically significant and positively signed causes of market leverage. These relationships largely match up with the documented empirical regularities (see, for instance Frank and Goyal[6]; p.195, "stylized fact 8"). In fact, they put these data relationships on a sounder methodological footing by validating them as causative effects, as well as showing that their economic effects are considerably larger than previously identified. This is particularly true regarding the role of firm size, where previous work has produced ambiguous evidence for a size effect. Here we show that firm size is strongly positively related to market leverage, and we use the causal framework to explain the earlier results.

One message of this manuscript is that empirical tests for capital structure theories should be based on sound causal models, not statistical hypotheses alone. Specifically, we argue that considerations about causal structures should largely replace considerations about goodness of fit in the model section process. In this context, we have shown how to use structural causal modeling to achieve these goals. There is a further benefit of the SCM approach, namely that it requires an explicit view for how data are generated by the theoretical model under consideration. One aim of this paper is to provide and test a causal framework for firms' capital structure decisions as a path towards future causal modeling in capital structure research.

Competing financial interests

The author declare no competing financial interests. SC received a part-time salary for this work. SK is an Operating Partner and Head of Research for Blackstone Credit's Systematic Strategies unit. The thoughts and opinions expressed in this article are those of the authors alone, and no other person or institution has any control over its content.

Authors contributions

SC and SK designed the study. SC performed the analysis. SC and SK wrote the paper.

Data availability

We used proprietary data and data from third parties (COMPUSTAT). In the main text we have provided the necessary information to reconstruct our dataset.

Acknowledgments

The authors would like to thank the Editor and two anonymous reviewers for their helpful comments on the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jfds.2022.09.002.

References

1. Modigliani F, Miller M. The cost of capital, corporation finance and the theory of investment. *American Economical Review*. 1958;48(3):261–297.
2. Kraus A, Litzenberger RH. A state-reference model of optimal financial leverage. *Journal of Finance*. 1973;28(4):911–922.
3. Myers SC, Majluf NS. Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economy*. 1984;13(2):187–221.
4. Baker M, Wurgler J. Market timing and capital structure. *Journal of Finance*. 2002;57:1–32.
5. Welch I. Capital structure and stock returns. *Journal Political Economy*. 2004;112:106–131.
6. Frank MZ, Goyal VK. Chapter 12 - trade-off and pecking order theories of debt. In: Eckbo BE, ed. *Handbook of Empirical Corporate Finance*. Handbooks in Finance.
7. Rajan RG, Zingales L. What do we know about capital structure? Some evidence from international data. *Journal of Finance*. 1995;50(5):1421–1460.
8. Wright PG. *The Tariff on Animal and Vegetable Oils*. Macmillan Company; 1928.
9. Haavelmo T. The statistical implications of a system of simultaneous equations. *Econometrica*. 1943;11(1):1–12.
10. Tinbergen J, Hendry DF, Morgan MS. *Determination and Interpretation of Supply Curves: An Examplep*. vols. 233–246. Cambridge University Press; 1995.
11. Rubin D. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Eduction Psychology*. 1974;66(5):688–701.
12. Imbens G. *Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics*. National Bureau of Economic Research; 2019. Working Paper 26104.
13. Hahn J, Todd P, Van der Klaauw W. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*. 2001;69(1):201–209.

14. Rosenbaum PR, Rubin B, Donald. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41–55.
15. Imbens G. Nonparametric estimation of average treatment effects under exogeneity: a review. *Review of Economics and Statistics*. 2004;86(1):4–29.
16. Wager S, Athey S. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of American Statistical Association*. 2018;113(523):1228–1242.
17. Splawa-Neyman J, Dabrowska DM, Speed TP. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statistical Science*. 1990;5(4):465–472.
18. Rubin DB. *Matched Sampling for Causal Effects*. Cambridge University Press; 2006.
19. Pearl J. Causal diagrams for empirical research. *Biometrika*. 1995a;82(4):669–688.
20. Pearl J, Glymour M, Jewell N. *Causal Inference in Statistics: A Primer*. Wiley; 2016.
21. Titman S, Wessels R. The determinants of capital structure choice. *Journal of Finance*. 1988;43(1):1–19.
22. Fama EF, French KR. Testing trade-off and pecking order predictions about dividends and debt. *Review of Financial Studies*. 2015;15(1):1–33.
23. Faulkender M, Petersen MA. Does the source of capital affect capital structure? *Review of Financial Studies*. 2005;19(1):45–79.
24. Hovakimian A, Kayhan A, Titman S. Credit rating targets. *SSRN Electronic Journal*. 2009. https://doi.org/10.2139/ssrn.1098351.
25. Hovakimian A, Li G. In search of conclusive evidence: how to test for adjustment to target capital structure. *Journal of Corporate Finance*. 2011;17(1):33–44.
26. Bessler DA, Yang J. The structure of interdependence in international stock markets. *Jounal of International Money Finance*. 2003;22(2):261–287.
27. Haigh MS, Bessler DA. Causality and price discovery: an application of directed acyclic graphs. *J Bus*. 2004;77(4):1099–1121.
28. Gow ID, Larcker DF, Reiss PC. Causal inference in accounting research. *Journal of Accounting Research*. 2016;54(2):477–523.
29. Su Z, Liu P, Fang T. Uncertainty matters in US financial information spillovers: evidence from a directed acyclic graph approach. *The Quarterly Review of Economics and Finance*. 2022;84:229–242.
30. Graham JR, Leary MT. A review of empirical capital structure research and directions for the future. *Annual Review of Financial Economics*. 2011;3(1):309–345.
31. Hovakimian A, Kayhan A, Titman S. Are corporate default probabilities consistent with the static trade-off theory? *Rev Financ Stud*. 2011;25(2):315–340.
32. Frank MZ, Goyal VK. Capital structure decisions: which factors are reliably important? *Financial Management*. 2009;38(1):1–37.
33. Merton R. On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance*. 1974;29(2):449–470.
34. Kealhofer S. Quantifying credit risk I: default prediction. *Financial Analysis Journal*. 2003;59.
35. Crosbie P, Bohn J. Modeling default risk. In: *World Scientific Reference on Contingent Claims Analysis in Corporate Finance*. World Scientific; 2019:471–506.
36. Gretton A, Fukumizu K, Teo CH, Song L, Schölkopf B, Smola AJ. A kernel statistical test of independence. In: *Proceedings of the 20th International Conference on Neural Information Processing Systems, NIPS'07*. Red Hook, NY, USA: Curran Associates Inc; 2007:585–592.
37. Zhang K, Peters J, Janzing D, Schölkopf B. Kernel-based conditional independence test and application in causal discovery. In: *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence, UAI'11*. Arlington, Virginia, USA: AUAI Press; 2011:804–813.
38. Strobl EV, Zhang K, Visweswaran S. Approximate kernel-based conditional independence tests for fast non-parametric causal discovery. *Journal of Causal Inference*. 2019;7(1), 20180017.
39. Leary MT, Roberts M. Do firms rebalance their capital structures? *Journal of Finance*. 2005;60(6):2575–2619.
40. Hovakimian A, Hovakimian G, Tehranian H. Determinants of target capital structure: the case of dual debt and equity issues. *Journal of Financial Economy*. 2004;71(3):517–540.
41. Chen L, Zhao X. On the relation between the market-to-book ratio, growth opportunity, and leverage ratio. *Finance Research Letters*. 2006;3(4):253–266.
42. Vuolteenaho T. What drives firm-level stock returns? *Journal of Finance*. 2002;57(1):233–264.
43. Boeken PA, Mooij JM. *A Bayesian Nonparametric Conditional Two-Sample Test with an Application to Local Causal Discovery*. 2021.
44. Runge J. Causal network reconstruction from time series: from theoretical assumptions to practical estimation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*. 2018;28(7), 075310.