Cash flow shocks and corporate liquidity*

Sebastian Gryglewicz† Loriano Mancini‡ Erwan Morellec§
Enrique Schroth¶ Philip Valta‖

September 18, 2019

Abstract

Theory has recently shown that corporate policies should respond differently to permanent or transitory cash flow shocks. We devise a novel filter to decompose cash flow shocks into permanent and transitory components. The policy choices of large publicly traded U.S. firms, such as cash holdings, credit line usage, and equity issuance, are related to the characteristics of the shocks estimated by our filter, i.e., volatilities, correlation and drift rates of the permanent and transitory shocks, as predicted by theory. Moreover, the interaction between the permanent and transitory cash flow shocks is strongly related to a firms leadership status within its industry.

Keywords: Cash flow risk, permanent and transitory shocks, liquidity management.
JEL Classification Numbers: G31, G32, G35.

*We thank Bruno Biais, Sudipto Dasgupta, Laurent Frésard, Patrick Gagliardini, Nicolae Gârleanu, Bart Lambrecht, Boris Nikolov, Gordon Phillips, Jean-Charles Rochet, Ailsa Röell, Clemens Sialm, Pascal St-Amour, Peter Swan, Elu von Thadden, Alexander Wagner, Josef Zechner and seminar participants at EDHEC, USI Lugano, University of Oxford, University of New South Wales, University of Illinois Urbana-Champaign, DePaul University, Warwick Business School, the CEPR Second Annual Spring Symposium in Financial Economics, European Finance Association Meetings, and the Swiss Finance Institute Research Days for comments. Loriano Mancini and Erwan Morellec acknowledge financial support from the Swiss Finance Institute.

†Erasmus University Rotterdam. Email: gryglewicz@ese.eur.nl.

‡USI Lugano and Swiss Finance Institute. E-mail: loriano.mancini@usi.ch.

§EPF Lausanne, Swiss Finance Institute, and CEPR. E-mail: erwan.morellec@epfl.ch.

¶EDHEC Business School and CEPR. E-mail: enrique.schroth@edhec.edu.

‖University of Bern. E-mail: philip.valta@ifm.unibe.ch.

Starting with Gorbenko and Strebulaev (2010) and DeMarzo, Fishman, He, and Wang (2012), a growing theoretical literature in corporate finance argues that corporate policies should respond differently to permanent and transitory cash flow shocks. Indeed, permanent shocks affect not only a firm's immediate productivity and cash flows but also its future productivity and cash flows. By contrast, while transitory shocks affect immediate cash flows, they are uninformative about future expected profitability.[1] While the decomposition of shocks between transitory and permanent components has been used productively over the years in many areas of economics, it has received little attention in empirical corporate finance.[2] This is relatively surprising given that the theoretical literature has shown that cash flow risk and cash flow shocks are key drivers of corporate policies.

The objective of this article is to start filling the existing void by empirically investigating whether permanent and transitory cash flow shocks are related to firms' liquidity decisions as predicted by theory. To do so, we start by decomposing the cash flow shocks of large publicly traded U.S. firms for the period 1971–2016 into permanent and transitory components. We then investigate whether the policy choices of these firms—cash holdings decisions, credit line usage, and equity issues—are related to the characteristics of the shocks they face, i.e., volatilities, correlation and drift rates of the permanent and transitory shocks. Our empirical analysis supports many of the predictions of recent dynamic corporate finance models analyzing the effects of permanent and transitory shocks on liquidity management. It also demonstrates the importance of accounting for the characteristics of both permanent and transitory shocks when explaining cross-sectional variation in corporate policies.

[1] Many types of firm or market shocks are transitory and do not affect long-term prospects. Examples include temporary changes in demand, delays in costumer payments, machine breakdowns, or supply chain disruptions. But long-term cash flows also change over time due to various firm, industry, or macroeconomic shocks that are of permanent nature. Examples include changes in technology, reductions of trade barriers, or changes in consumer preferences.

[2] A number of asset pricing papers (see, e.g., Cochrane (1994), Cohen, Gompers, and Vuolteenaho (2002), Bansal, Dittmar, and Kiku (2008), Garleanu, Kogan, and Panageas (2012a), or Garleanu, Panageas, and Yu (2012b)) use such a decomposition to analyze stock returns and risk premia on stocks. This decomposition is also used in market microstructure to analyze price efficiency (see, e.g., Glosten and Harris (1988), Brennan and Subrahmanyam (1996), or Boehmer and Wu (2013)). The literature on income processes also often seeks to decompose shocks into permanent and transitory components; see, e.g., Blundell, Pistaferri, and Preston (2008), Meghir and Pistaferri (2004), or Gottschalk and Moffitt (2009). The decomposition of income shocks between permanent and transitory components has found interesting applications in the life-cycle portfolio choice literature; see, e.g., Cocco, Gomes, and Maenhout (2005). In the time series literature, the permanent-transitory model is known as the unobserved component decomposition, in which the permanent part is the trend and the transitory component is named the cyclical innovation; see Hamilton (1994, Chapter 17).

We begin our empirical analysis by testing whether firm cash flow shocks include a permanent, non-stationary, component. To this end, we implement two tests, the Augmented Dickey–Fuller (ADF) test and the Kwiatkowski, Phillips, Schmidt, and Shin (1992) (KPSS) test. Using both tests, we find large evidence in favor of a permanent component in firms' cash flow shocks. Having established that corporate cash flows include a permanent, non-stationary, component, we then turn to examining the relation between the characteristics of cash flow shocks and liquidity choices. Assessing the importance of permanent and transitory shocks for corporate policies faces two challenges. First, we need to be able to identify permanent shocks separately from transitory shocks. Second, we need to obtain reliable estimates of their characteristics, in order to relate these to corporate policies.

To address these challenges, we adopt the cash flow structure put forth by Décamps, Gryglewicz, Morellec, and Villeneuve (2017), which nests as special cases many of the cash flow environments used in dynamic liquidity management, agency, capital structure, and real-options models. In this model, cash flows are subject to profitability shocks that are permanent in nature. In addition, for any given level of profitability, cash flows are also subject to short-term shocks that may be purely transitory but may also be correlated with permanent shocks. To identify potentially correlated permanent and short-term shocks, we develop a novel Kalman filtering technique that is consistent with this theoretical structure. We then use our filtering technique and maximum likelihood to estimate from panel data on actual firm cash flows the parameters driving the cash flow shocks—that is, the growth rates and volatilities of permanent and short-term shocks as well as their correlation—that best explain observed firm cash flows.

To determine whether our estimates of the characteristics of cash flow shocks are meaningful, we proceed with a major out-of-sample exercise. Specifically, we show that the estimated cash flow shock characteristics imply asset volatilities that match the actual asset volatilities of the firms in our sample. Importantly, actual asset volatilities are not used in the estimation of cash flow characteristics. Equally important, cash flow characteristics have been estimated using cash flow data from operating earnings, without imposing any model restriction about corporate policies. We find a close match between implied and empirical asset volatilities. To the best of our knowledge, ours is the first study to reproduce actual asset volatilities using cash flow-based volatility measures.

We proceed by empirically investigating the relation between corporate cash holdings and cash flow characteristics. Consistent with theory, we find that firms with a higher growth rate of permanent shocks and a higher uncertainty arising from permanent and transitory shocks have larger cash holdings. A one standard deviation increase in the volatility of permanent or transitory shocks is associated with an average increase of 8% to 14% in the ratio of cash to total assets. A unique prediction of Décamps et al. (2017) is that firms with more highly correlated permanent and short-term shocks hold less cash. Accordingly, a one standard deviation increase in this correlation coefficient is associated with an average decrease of 4.7% of the cash-to-asset ratio. Our empirical analysis also shows that our estimated correlation and volatilities subsume an important part of the previously documented relation between industry cash flow risk and cash holdings. That is, previous studies had focused on cash flow uncertainty as one of the important determinants of cash holdings (e.g., Bates, Kahle, and Stulz (2009)). Our analysis demonstrates that cash holdings decisions can be better understood as arising from *two* separate types of cash flow shocks: permanent and transitory.

Because equity issues are partly used to replenish cash reserves, as empirically shown by Kim and Weisbach (2008) or McLean (2011), we next examine the relation between the amount of equity issued and cash flow characteristics. As predicted by theory, we find that firms issue more equity when short-term and permanent cash flow shock volatilities are high and when the growth rate of permanent shocks is high. In addition, firms issue less equity when the correlation between short-term and permanent shocks is high. Like cash, credit lines are an important source of liquidity to guard against negative cash flow shocks and to give firms the option to exploit growth opportunities. As with cash, theory therefore predicts that the use of credit lines is negatively related to the correlation between permanent and short-term cash flow shocks, and positively related to the volatilities of both shocks and to the growth rate of permanent shocks. We find again strong support for these predictions in the data. Notably, a one standard deviation increase in the correlation between permanent and short-term shocks is associated with a decrease in credit lines usage of 10.3% relative to the sample mean. Altogether, the evidence supports the unique feature of models with permanent and transitory shocks that a higher correlation between permanent and short-term shocks actually implies *less* risk, as firms expect to generate cash flows when needed after positive permanent shocks.

We next examine the adjustment of cash balances towards the cash target. Theory predicts that the cash flow sensitivity of savings should be positive when the correlation between permanent and short-term cash flow shocks is positive, and negative otherwise. In addition, the absolute sensitivity should increase with the ratio of the volatility of permanent shocks to the volatility of short-term shocks. In line with these predictions, we find that the cash flow sensitivity of savings switches signs depending on the shock correlation when the ratio of volatilities is relatively high, which is precisely where the absolute cash flow sensitivity of savings is expected to be highest.

We conclude our analysis by offering a deeper interpretation of the correlation between permanent and short-term shocks. This correlation uncovers a novel feature of the cash flow data and our approach offers a rich cross sectional variation in its estimates. We show that the correlation is not a characteristic intrinsic to each industry but rather a feature specific to firms within industries. The correlation is strongly related to how firms perform within their industries. Indeed, as shown by theory, a positive correlation allows firms to better exploit and manage cash flow shocks. In the data, very few firms exhibit such positive correlation in any given industry. These firms are leaders in their industry and high shock correlation appears to be associated with systematic high performance and innovation leadership. This pattern is present across all industries. Our findings are consistent with the interpretation of the shock correlation as a reduced-form measure of a first-mover advantage.

Our paper relates to the growing theoretical literature on the effects of permanent and transitory shocks on corporate policies. DeMarzo et al. (2012), Hoffmann and Pfeil (2010), Hackbarth, Rivera, and Wong (2018), and, Gryglewicz, Mayer, and Morellec (2019) examine the effects of permanent and transitory cash flow shocks on optimal compensation and investment in dynamic moral hazard models. Gorbenko and Strebulaev (2010) and Hackbarth, Miao, and Morellec (2006) develop a dynamic capital structure model in which cash flows are subject not only to permanent shocks as in Leland (1998) but also to temporary (but persistent) shocks. Grenadier and Malenko (2010) build a real options model in which firms are uncertain about the persistence of past shocks and have the option to learn before investing in a single irreversible investment opportunity. Décamps et al. (2017) examine the effects of permanent and transitory shocks on cash holdings, credit lines usage, equity issues, and risk management in a model with financing frictions.

As relevant as it is to analyze the effects of transitory and permanent shocks on corporate policies, there are surprisingly few studies in the empirical literature addressing this problem. In an early study, Guay and Harford (2000) show that firms that choose dividend increases also exhibit persistent cash flow shocks but those that use repurchases exhibit relatively more cash flow reversion. Chang, Dasgupta, Wong, and Yao (2014) show that the relatively more financially constrained firms allocate more of their transitory cash flow to savings than investment. Lee and Rui (2007) find that share repurchases are used to pay out cash flows that are potentially transitory, thus preserving financial flexibility relative to dividends. Guiso, Pistaferri, and Schivardi (2005) examine the allocation of risk between firms and their workers and show that firms absorb transitory shocks fully but insure workers against permanent shocks only partially. Lastly, Byun, Polkovnichenko, and Rebello (2018) examine the separate effects of persistent and transitory cash flow shocks on leverage decisions.

Our paper advances this literature in three important ways. First, we establish the presence of permanent components for the majority of Compustat firms' cash flows and develop a novel Kalman filtering technique tailored to identify permanent and short-term shocks within the canonical cash flow model used in dynamic corporate finance. Our method allows for an arbitrary correlation between these shocks and for an unprecedentedly high degree of granularity in the estimation of cash flow characteristics. This filter performs much better empirically than the restrictive "off-the-shelf" filters used in all prior studies, leading to a sharper and more robust analysis of the relation between cash flow characteristics and corporate policies.[3] Second, we demonstrate that the distinction between transitory and permanent shocks is relevant for the larger set of policies that are of interest to financial economists, namely financing, cash holdings, and payout policies. Third, we show that firm

[3]Prior studies use the Hodrick–Prescott filter or the Beveridge–Nelson decomposition to separate a time series into a trend (persistent) component and a cyclical (transitory) component. The use of these filters is potentially problematic because they cannot handle missing values (which are pervasive in large cash flow panels), they are not suited to study correlations between persistent and short-term shocks (which are an important feature of cash flow data and affect policies), and they are not designed to recover volatilities of permanent shocks.

policies are strongly associated with the interaction (correlation) between short-term and permanent shocks and find that this interaction is closely related to the firm's performance and leadership status within its industry.

1 Corporate cash flows data

This section studies the time series properties of the cash flow data. In order to understand the nature of shocks affecting operating cash flows at the firm level, as well as determining which class of models best describe their dynamics, we start by investigating whether cash flows are stationary.

1.1 Sample

We collect accounting data from Compustat between 1971 and 2016, stock price data from CRSP, and credit line data from Thomson Reuters Dealscan. We exclude financial services firms (SIC codes 6000 to 6999), utilities (SIC codes 4900 to 4999), and other regulated firms (SIC codes 8000 to 9999). We convert all data into 2000 constant dollars using the GDP deflator. We drop firm-years with annual asset growth rates above 500%.[4] Finally, we winsorize the firm-level variables at the 1st and 99th percentiles. Our final sample includes 186,769 firm-years for 9,232 firms.

1.2 The operating cash flows variable

We define *Operating cash flows* as EBITDA minus changes in working capital, where the definition of working capital follows Chang et al. (2014). These cash flows thus represent the component of the firm's overall cash flow that is derived from its operations and short-term investment but not from its financing decisions, such as cash holdings, equity issuance, dividends, or credit line usage. This definition of operating cash flow is the same as in, e.g., Décamps et al. (2017).

Insert Table 1 Here

[4] This screen intends to eliminate firms that experience major corporate events such as a merger or a restructuring (see, e.g., Almeida, Campello, and Weisbach (2004)).

To make firms of different sizes comparable, we divide each year's operating cash flow by the firm's *initial* value of total assets. This normalization does not affect the time series properties of operating cash flows because the initial value of assets is constant over time. Table 1 contains the definitions and descriptive statistics of our operating cash flow variable as well as other firm-specific characteristics that we use in the empirical analysis.

1.3 Time series properties of operating cash flows

We run two tests to determine whether firm cash flows include a permanent (non-stationary) component or are driven by purely transitory shocks: The Augmented Dickey–Fuller (ADF) test and the Kwiatkowski et al. (1992) (KPSS) test. In the ADF test, the null hypothesis is that operating cash flows are non-stationary and have a unit root, while the alternative hypothesis is that they are stationary and follow an autoregressive process with a drift. Given the risk of failing to reject the null hypothesis with small sample sizes, we also implement the KPSS test, in which the null hypothesis is that cash flows are stationary while the alternative is that they follow an integrated process. In KPSS tests, the bias with small samples is towards not rejecting stationarity. We run the ADF and KPSS tests for each firm's cash flow time series in our panel. Table 2 summarizes the results.

Insert Table 2 Here

The first row in Panel A of Table 2 shows the results for the ADF and KPSS tests over the 332 firms in our sample that have no missing cash flow observations for the full sample period. For 97% of the firms in this subsample, in which both tests have the highest power, the ADF test does *not* reject the null hypothesis of non-stationarity at the 1% level. For 91.9% of these firms, the KPSS tests *reject* the null hypothesis of stationary cash flows at the 10% level. The high rejection rate is remarkable for the KPSS test, given that rejecting the hypothesis of stationarity in relatively short samples is typically rare.

The ADF and KPSS tests results also strongly suggest non-stationarity for firms with fewer cash flow observations: For 82.7% and 84.7% of all 1,349 firms with at most one third of cash flow observations missing, the ADF tests do not reject non-stationarity and the KPSS tests reject stationarity at the 10% level, respectively. Even as the power decreases further, the KPSS tests still reject the hypothesis of stationarity at the 10% level for 70.3% of the 5,819 firms with at most two thirds of missing cash flow observations.[5] We also find considerable evidence of non-stationarity for the ratio of cash flows to (one year) lagged assets (*Cash flow-to-assets*),which is often used in empirical corporate finance research. Panel B in Table 2 shows high proportions of firms for which non-stationarity cannot be rejected under the ADF test and stationarity is rejected under the KPSS test.

In sum there is strong evidence that a majority of firms' cash flows include a permanent component. Since optimal corporate policies will respond differently to permanent and transitory cash flow shocks, models of the firm's operating cash flows must include a permanent component to describe observed policies accurately. We also detect a permanent component for most cash flow-to-assets ratios, implying, for example, that an autoregressive process is not a good description of a firm's ratio of cash flow to lagged assets. Indeed, shocks to an autoregressive component do not have a permanent but at best persistent impact on future cash flows and in practice die out quickly.

Finally, a pronounced feature in the cash flow data is the pervasiveness of missing observations: 56% of the full panel of Compustat firm-year observations are missing. Therefore, any sound econometric analysis of cash flow data needs to take missing values into account.

2 Estimation of the operating cash flows model

Having established that corporate cash flows are exposed to permanent shocks, we now decompose cash flow shocks into permanent and transitory components and estimate their primitive parameters. In the following, we estimate a discrete time version of the continuous time cash flow model in Décamps et al. (2017), in which cash flows are driven by permanent and transitory shocks (see Appendix A). This is without loss in generality as this model nests as special cases the cash flow environments used in real-options models and dynamic capital structure models, in which cash flow shocks are permanent (see, e.g., Abel and Eberly (1994) or Leland (1998)), and the cash flow environments used in liquidity management models and

[5]The critical values for the ADF and KPSS tests cannot be computed for firms with more than two thirds of the time series observations missing.

dynamic agency models, in which cash flow shocks are purely transitory (see, e.g., Bolton, Chen, and Wang (2011) or DeMarzo and Sannikov (2006)).

2.1 The model

A standard assumption in corporate finance is to assume that firm cash flows follow a Gordon growth model and are only subject to permanent shocks. Our cash flow model builds on this assumption but additionally allows cash flows to be subject to transitory shocks. Specifically, the cash flow model in state space form consists of the following transition and measurement equations:

$$P_t = (1+\mu)\, P_{t-1} + \sigma_P\, P_{t-1}\, \varepsilon_t^P \tag{1}$$

$$A_{i,t} = P_t + \sigma_A\, P_{t-1}\, \varepsilon_{i,t}^A \tag{2}$$

where P_t is the unobserved asset productivity with constant growth rate μ and permanent volatility $\sigma_P > 0$, and $A_{i,t}$ is the operating cash flow of firm i, $i = 1, \ldots, N$, in year t, with short-term volatility $\sigma_A > 0$. The permanent shock ε_t^P influences cash flows permanently by affecting the productivity of assets. The short-term shock $\varepsilon_{i,t}^A$ impacts the cash flow directly and may affect the firm's long-term prospects. We allow short-term and permanent shocks to be correlated with correlation coefficient $\rho \in (-1, 1)$. Hence, the short-term shock can be written as

$$\varepsilon_{i,t}^A = \rho \varepsilon_t^P + \sqrt{1-\rho^2}\varepsilon_{i,t}^T \tag{3}$$

where $\varepsilon_{i,t}^T$ is a purely transitory shock uncorrelated with ε_t^P. Both ε_t^P and $\varepsilon_{i,t}^T$ are distributed as $\mathcal{N}(0,1)$. When $\sigma_A = 0$, cash flows follow a Gordon growth model and are only subject to permanent shocks. When $\mu = \sigma_P = 0$, cash flow shocks are i.i.d. and follow a purely stationary process. For $\sigma_A > 0$ and $\sigma_P > 0$, our specification implies a rich cash flow structure that allows for permanent and transitory cash flow shocks.

The model in equations (1)–(2) captures the heterogeneity of firms' operating cash flow exposures to long-term and short-term risk via different combinations of values of the parameters σ_P, σ_A and ρ. Estimation of these parameters at the highest possible level of granularity allows us to test the rich empirical predictions that different combinations have on observed financing policies. As shown by Décamps et al. (2017), the parameters μ and ρ in (1) and (3) have unique predictions on liquidity and financing policies.

2.2 Estimation

The goal of our estimation is to separately identify permanent and short-term shocks and estimate the cash flow parameters (μ, σ_A, σ_P, and ρ). If short-term and permanent shocks were uncorrelated, the model in (1)–(2) could be estimated with a standard Kalman filter using Maximum Likelihood. Given model parameters, the Kalman filter would recover the unobserved asset productivity P_t, which would enter the likelihood function of observed cash flows. However, because the shocks are correlated, the standard Kalman filter is biased and inconsistent and, therefore, cannot be used. This problem is reminiscent of an endogeneity issue in regression analysis, with the major difference that the regressor—the asset productivity—is unobserved and needs to be filtered out.

We solve this problem by theoretically regressing $\sigma_A \, P_{t-1} \, \varepsilon^A_{i,t}$ on $\sigma_P \, P_{t-1} \, \varepsilon^P_t$, and by properly transforming the measurement equation (2). Because of this transformation, we need to derive a novel Kalman filter. Appendix B describes our transformed Kalman filter and the estimation method in detail. We note that our method does not involve any approximation of the model in (1)–(2). In fact, our algorithm is a generalized Kalman filter that reduces to the standard version when $\rho = 0$.

Our generalized filter differs fundamentally from other very popular methods to separate a time series into a trend (persistent) component and a cyclical (transitory) component, namely, the Hodrick–Prescott (HP) filter and the Beveridge–Nelson (BN) decomposition.[6] First, neither method is suited to study correlations between persistent and short-term shocks, as the HP filter presumes no correlation between trend and cyclical components, whereas the BN decomposition imposes a perfectly negative correlation between trend and cyclical shocks.[7] Second, neither is specifically designed to recover volatilities of permanent

[6] In the empirical corporate finance literature, the Hodrick–Prescott filter has been applied for instance by Byun et al. (2018) and the Beveridge–Nelson decomposition by Chang et al. (2014).

[7] The Beveridge–Nelson decomposition would rely on an ARMA(2,2) model for the changes of cash flows.

shocks. This aspect is problematic because permanent shocks are a major driver of cash flows, as we show with formal tests in Section 1. Failing to recover properly the volatilities of permanent shocks introduces biases when estimating volatilities of transitory shocks. Finally, neither can handle missing values. Indeed, such methods were developed to decompose typically complete annual aggregate time series such as consumption or GDP. Kalman filtering is by design a method of estimation and imputation of missing data (see Appendix B.3). Since only 332 firms, out of 9,232 firms in our panel, have complete time series of cash flows, applying HP or BN forces the researcher to drop the vast majority of the firms in Compustat.

A Monte Carlo analysis, described in detail in Appendix C, confirms that our generalized filter performs better in Monte Carlo simulation than the traditional Kalman filter, in which $\rho = 0$. Not surprisingly, given that it is tailored specifically to estimate the cash flow model with correlated permanent and short-term shocks, our filter also outperforms the HP filter and the BN decomposition (see Appendix C).

2.3 Identifying assumptions

If the shock correlation ρ were equal to zero, the cash flow model in (1)–(2) would be a classic state space model with Gaussian likelihood. The remaining three parameters would be identified from the unique global maximizer of the likelihood function (see Appendix B for details). To identify the shock correlation, this model assumes that transitory shocks are firm-year specific while permanent shocks have a "systematic" nature, i.e., that they are common to all firms in a set of equal asset productivity. Indeed, correlated permanent and short-term shocks are not identified if both are firm-specific.

To provide an intuition for this identifying assumption, consider a conventional model of asset productivity in which an observable process is driven by the sum of persistent shocks (modelled as an AR(1) process) and short-term shocks (modelled as a white noise process). When both shocks are unobservable, the challenge of identifying their correlation is similar to identifying the correlation ρ in (1)–(2). As we show formally in Appendix B.4, the

A well-known limitation of the BN decomposition is that the trend component is always mechanically more volatile than the original time series. It is so because, by design of the filter, trend shocks are perfectly negatively correlated with cyclical shocks. Thus, the original time series (sum of trend and cyclical components) is always less volatile than the trend component.

correlation between firm-specific persistent and short-term shocks is not identified because this parameter enters the autocovariance function of the individual firm productivity only as a multiplicative constant to both persistent and short-term volatilities.[8] If instead, persistent shocks are common across firms, as per our identifying assumption, then the time series of the cross-sectional average productivity provides additional and non-redundant moment conditions to identify the shock correlation. In essence, the common persistent shock has the interpretation of a "systematic" cash flow factor, as in the classic factor models that are routinely estimated in the empirical asset pricing literature.

The implication of this result is that the firm-by-firm estimation of the model in (1)–(2) is infeasible. Therefore, we estimate the model parameters for groups of firms, where firms of the same group are exposed to the same permanent shocks. We form the smallest groups possible to achieve maximum granularity across our sample firms.

2.4 Grouping firms with similar cash flow dynamics

Our goal is to propose and empirically validate a specific structure of permanent and short-term cash flow shocks. To achieve this goal, we must estimate the parameters of the cash flow model with the highest possible level of precision and granularity to show that differences in these estimates explain observed cross-sectional heterogeneity in corporate policies, as predicted by theory.

We estimate the cash flow model in (1)–(2) for each of *many* small groups of firms. We assume that all firms within each group are homogenous in that they have the same parameters $\mu, \sigma_P, \sigma_A, \rho$ and asset productivity, P_t. Fitting the model to relatively small samples allows us to achieve greater estimation accuracy because the model parameters can adjust to the data features of each specific group of firms. Moreover, we obtain a large set of estimates of the cash flow model's deep parameters, as opposed to just one or a few sets for the representative firms. Since the limiting case—which is to estimate the cash flow model firm-by-firm—is not feasible, estimation by small groups maximizes the cross sectional variation in these estimates and enables a direct test of the link between deep parameter

[8] A proof that the shock correlation in (1)–(2) is not identified when permanent shocks are firm-specific is beyond the scope of this paper given that the model features a multiplicative, non-stationary process.

heterogeneity and predicted corporate policies.

To group firms, we adopt two sequential criteria that are motivated by the assumption that permanent shocks are common to all firms in the group. The first is the three-digit SIC industry code. We expect firms within the same three-digit SIC industry to be exposed to similar short-term volatility (e.g., industry demand uncertainty) and similar permanent shocks (e.g., technology or regulatory shocks). The second is the firm's cash flow growth rate: Within each three-digit SIC industry, we group firms based on their average annual growth rate of cash flows. In the long-run, firms with similar asset productivity will have similar average cash flow growth rates. For the precision of our parameter estimates, based on the simulation evidence in Appendix C, we impose the additional requirement that each group includes at least 10 firms. Because the number of firms in any given industry is not generally a multiple of 10, the last group of firms in each three-digit SIC code will include between 10 and 19 firms.

Applying the criteria above, our sample of 9,232 firms is split into 794 SIC3-cash flow growth groups. As an example, Figure 1 shows the cash flows of one group of firms in the 100 SIC code. Missing observations in firm cash flows are evident in the interrupted time series of firm cash flows. Because our groups are relatively small, our parameter estimates can potentially exhibit substantial variation even within three-digit SIC industries.

Insert Figure 1 Here

To assess the homogeneity of firms within each group, we decompose the total variation of several firm-specific outcome and policy variables into the between- and within-group components. For each characteristic, we compare the similarity within and heterogeneity between our estimation groups to those implied by other narrow industrial classifications. Table 3 shows that, relative to the four-digit SIC or the 38 Fama and French (1997) industries, our classification produces less *within-group* variation for the ratios of annual sales-to-assets, earnings-to-assets and average sales growth, as well as for key policy variables such as cash holdings, the rates of savings and equity issuance, the size of the credit lines (relative to total debt), the capex-to-assets and debt-to-assets ratios. Our grouping also implies more within-similarity in the ratio of R&D expenditure to sales, the number of patents and their market value, according to Kogan, Papanikolaou, Seru, and Stoffman (2017). Remarkably, grouping only by long-run similarity in the average cash flow growth rate within each three-digit SIC industry produces similarities across many other dimensions.

Insert Table 3 Here

Table 3 shows that our grouping method also produces the most variation *between-group* for as many firm characteristics relative to the four-digit SIC or the 38 Fama and French (1997) industrial classifications. The only exception is the number (but not market value) of patents.

3 Cash flow model estimation: Results

3.1 Parameter estimates

Table 4 summarizes the Maximum Likelihood (ML) estimates of the model's four parameters, μ, σ_P, σ_A, and ρ, for all the 794 three-digit SIC-cash flow growth rate groups (Panel A), their precision (Panel B), and the correlation between the parameter estimates (Panel C). We winsorize the estimates at the 1st and 99th percentiles when they approach their respective lower and upper bounds (i.e., near -1 and 1 for the shock correlation, and near the zero value for each volatility), and at 10th and 90th percentiles otherwise.

Insert Table 4 Here

The estimates of the correlation between permanent and short-term shocks, ρ, exhibit significant variation across groups. Panel A shows that the 5th percentile of the estimated correlations is -0.25 and the 95th percentile is 0.27 (with a minimum of -0.94 and a maximum of 0.93, unreported). The average estimated correlation is -0.052, with almost 80% of the estimates being negative. Panel B shows that close to 60% of the estimated ρ are significantly different from zero with 95% confidence.

The estimates of μ exhibit an interquartile range from 6.6% to 23.7%, with an average productivity growth of 16.4%. Almost 80% of these estimates are significantly different from zero with 95% confidence. These estimates show that our grouping procedure has effectively captured important differences in latent productivity growth rates.

The average estimated short-term volatility $\hat{\sigma}_A$ is larger than the average estimated permanent volatility $\hat{\sigma}_P$, although this result is driven by a few very large estimates of the former. Indeed, the permanent shock volatility has a higher median and varies less across groups. Note that these volatility estimates are not directly comparable to observed asset volatilities nor to estimates derived from calibrations of cash flow models with one shock only. We assess whether our estimated $\hat{\sigma}_P$, $\hat{\sigma}_A$, and $\hat{\rho}$ can reproduce actual cash flow volatilities as well as asset volatilities in Section 3.2 below.

Another important result in Table 4 is that the estimates of all four parameters exhibit significant within and between three-digit SIC industry variation. This result underscores the success of our grouping procedure in capturing substantial parameter heterogeneity within industries. As a consequence, our estimates have the potential to not only identify a relation between differences in cash flow parameters and policies across but also within industries.

Next, we carry out two exercises in order to gauge the economic accuracy of the parameter estimates in Table 4. The first is an out-of-sample validation exercise in which we compare the asset volatilities implied by our estimated cash flow model to the actual distribution of asset volatilities of Compustat firms. As noted by Gorbenko and Strebulaev (2010), an important limitation of standard EBIT models with only permanent shocks (e.g., Leland (1994)) is that asset growth volatilities are equal to cash flow volatilities. Hence, this exercise evaluates the extent to which our estimates help reconcile the large differences between relatively high cash flow volatilities and the much more moderate asset volatilities in the data. The second exercise is an in-sample analysis that compares the distributions of actual cash flow volatilities to the estimated volatilities of permanent shocks. This exercise asks whether our predicted asset volatilities, which are in line with the data, are driven by the modelling assumptions or instead by robust inference based only on cash flow data, using the model in (1)–(2).

3.2 Asset growth volatilities

To compute the model-implied asset volatilities, we employ the model of Décamps et al. (2017) presented in Appendix A. The model uses the cash flow process estimated above and solves for optimal financing and liquidity policies and firm value. It thus quantitatively links the cash flow parameters (μ, σ_A, σ_P, and ρ) to asset volatility.

We calculate asset volatilities at the group level, consistently with the level of granularity of the estimation of cash flow parameters. The empirical asset volatilities are estimated as weighted averages of equity and debt volatilities following the approach in Bharath and Shumway (2008) using daily stock returns. Model-implied volatilities are averages of all firm level volatilities within a group using the estimated cash flow parameters reported in Table 4 for each group of firms. We drop groups of firms with insufficient stock price data or where the model cannot be solved given parameter values, winsorizing actual and predicted asset volatilities at the 5th and 95th percentiles.[9]

Insert Table 5 Here

Table 5 presents a comparison of the empirical and model-implied asset volatilities. The average and median empirical asset volatilities are 0.424 and 0.392, respectively. In the baseline case, the model-implied volatilities are sightly higher, at 0.467 and 0.414 respectively, but very close to the empirical ones. It is remarkable that model-implied asset volatilities appear to match actual asset volatilities that were not used during the estimation process. While very similar at the center of the distribution, the model-implied volatilities tend to be somewhat more extreme in the tails compared to the empirical ones. This suggests that there could be some forces beyond those in the Décamps et al. (2017) model moderating the volatility of real firms' asset values. Additional rows for model-implied distributions in Table 5 present calculations based on departures from the baseline parameters. The results show that the similarity between the empirical and model-implied distributions is robust

[9]The remaining model parameters follow the choices in Décamps et al. (2017) and in Bolton, Chen, and Wang (2013): the risk-free rate $r = 0.08$, the carry cost of cash $\lambda = 0.02$, proportional equity issuance cost $p = 1.06$, fixed equity issuance cost $\Phi = 0.002$, the market price of risk of temporary and permanent shocks $\eta^T = \eta^P = 0.4$, and the correlation of temporary and permanent shocks with market shocks $\xi^T = \xi^P = 0.4$; cash holdings are assumed at the target level.

and driven by the estimates of the parameters of the cash flow dynamics rather than the assumed values for the remaining parameters.

3.3 Permanent shocks volatilities

Our cash flow model estimates imply reasonable asset volatilities, and the results are robust to different calibrations of the other model parameters. The mapping of very high cash flow volatilities into moderate asset volatilities is not the result of assumptions of the corporate finance model but of the inference of our cash flow model from real cash flow data. Indeed, in our sample, the median of the standard deviations of annual cash flow growth is a very high 210%, yet our cash flow model infers a median estimate of σ_P of 65%.

To understand why our model makes such inference, consider how σ_P is identified. Let $\overline{A}_t$ denote the cross-sectional average cash flow within each group of firms, i.e., $\overline{A}_t = \sum_{i=1}^{N} A_{i,t}/N$, for $t = 1, \ldots, T = 46$ years. Let R_t denote the relative change of $\overline{A}_t$. This observable rate is approximately equal to the relative change of the unobservable P_t (with equality if $N \to \infty$, and $\rho = 0$ or $\sigma_A = 0$), i.e.,

$$R_t \equiv \frac{\overline{A}_t - \overline{A}_{t-1}}{\overline{A}_{t-1}} \approx \frac{P_t - P_{t-1}}{P_{t-1}}. \tag{4}$$

To illustrate, Figure 1 shows $\overline{A}_t$ superimposed to firm cash flows for a select group of low-correlation firms (-0.07): The time series trajectory of $\overline{A}_t$ mimics the estimated asset productivity $\hat{P}_t$ for this group. Absent short-term shocks ($\varepsilon^A_{i,t} = 0$), the time series standard deviation of R_t would be approximately equal to σ_P, as $A_{i,t} = P_t$ in (2). Otherwise, the approximation in (4) implies that the estimated volatility of permanent shocks is expected to be positively related to, albeit smaller than, the time series standard deviation of R_t.

Insert Figure 2 Here

Figure 2 shows the scatter plot of the time series sample standard deviations of R_t against model-inferred volatilities of permanent shocks, $\hat{\sigma}_P$, for the 794 groups in our sample. As predicted, this figure shows a strong positive association between standard deviations of R_t and the estimates of σ_P, providing direct evidence that our estimates of σ_P are capturing the volatility of permanent shocks. Moreover, standard deviations of R_t are generally larger than the estimates of σ_P, meaning that short-term shocks are present in cash flows data.

4 Cash flow risk parameters and corporate policies

4.1 Empirical predictions

In this section, we test the main predictions of the Décamps, Gryglewicz, Morellec, and Villeneuve (2017) model, which subsume those in other dynamic models with financing frictions but without permanent shocks such as Décamps, Mariotti, Rochet, and Villeneuve (2011), Bolton, Chen, and Wang (2011), Gryglewicz (2011), Bolton et al. (2013), Hugonnier, Malamud, and Morellec (2015), or Malamud and Zucchi (2019). In the Décamps et al. (2017) model, firms are subject to both permanent and transitory cash flow shocks and face financing frictions. The potential losses associated with negative cash flow shocks can be covered using cash holdings or credit lines or by raising outside funds at a cost. Appendix A summarizes the model in Décamps et al. (2017).

Consider first the predictions of the model for optimal cash holdings and credit lines. The model implies that target cash holdings and the size of credit lines should increase with the volatilities of both transitory and permanent shocks as an increase in volatility increases the likelihood of losses and the expected cost of raising external capital. A second and more surprising prediction of the model is that cash holdings and the size of credit lines should decrease with the correlation between short-term and permanent shocks. This is not immediately expected because an increase in correlation between two transitory shocks would imply less diversification and more risk and therefore an increase in cash holdings and the size of credit lines. To understand this prediction of the model, note that long-term prospects and firm value improve following positive permanent shocks, making it optimal to hoard more cash. The firm therefore benefits from *increased* correlation between short-term and permanent shocks because it expects to generate cash flows when they are needed to increase cash holdings after positive permanent shocks. This mechanism implies that the firm requires smaller cash reserves or credit lines when the correlation between permanent and short-term cash flow shocks is larger. A third (and also unique) prediction of the model is that target cash holdings and the size of credit lines should increase with the growth rate of productivity, due to the increase in firm value and in the precautionary value of cash reserves and credit lines.

Because equity issues are partly used to replenish cash reserves—as empirically shown by Kim and Weisbach (2008) or McLean (2011)—the Décamps et al. (2017) model predicts that the total equity issued should decrease with the correlation between short-term and permanent shocks and increase with the volatility of both short-term and permanent shocks. An additional and unique prediction of the model is that more profitable firms should raise more funds when accessing financial markets: Hence, the total equity issued over a fixed period of time should increase with the growth rate of productivity.

Consider next the cash flow sensitivity of cash savings. The model predicts a positive sensitivity, as in Almeida, Campello, and Weisbach (2004), if short-term and permanent shocks are positively correlated, but a negative sensitivity, as in Riddick and Whited (2009), if this correlation is negative. The intuition for this result is that when permanent and short-term shocks are positively correlated, the sensitivity of cash holdings to cash flow shocks is driven by the positive relation between profitability and the marginal value of cash, which implies that the firm optimally retains a part of a positive cash flow shock if profitability increases. These predictions are in sharp contrast to those in dynamic models with financing frictions but without permanent shocks, in which firms retain all earnings if they are below their cash target or distribute all positive cash flows if they have reached it. That is, the predicted propensity to save from cash flows is either one or zero in these models.

To summarize, the testable hypotheses of the Décamps et al. (2017) model are:

HYPOTHESIS 1: *Cash holdings, the total equity issued per year, and the size of credit lines decrease with the correlation between short-term and permanent shocks, and increase with the volatilities of short-term and permanent shocks and the growth rate of permanent shocks.*

HYPOTHESIS 2: *The sign of the sensitivity of cash savings to cash flows equals the sign of the correlation between short-term and permanent shocks.*

4.2 Cash holdings

To test whether the correlation between permanent and short-term cash flow shocks is negatively related to cash holdings, we augment the standard cash holdings regression model, in which the dependent variable is the value of cash and marketable securities divided by the book value of total assets (see, e.g., Bates, Kahle, and Stulz (2009)), with the estimated shock correlations $\hat{\rho}$:

$$\textit{Cash holdings}_{i,t} = \beta_0 + \delta_t + \beta_\rho \times \hat{\rho} + \boldsymbol{\beta}_{\text{Controls}} \times \mathbf{Controls}_{i,t-1} + u_{i,t}. \tag{5}$$

We use the subscripts i for firms and t for years. Our estimate of ρ is constant over time but varies across the 794 groups used for the cash flow model estimation. As in Bates et al. (2009), we control for the *Industry cash flow volatility*, which varies yearly at the two-digit SIC industry, for firms' growth opportunities with the lagged market-to-book ratio (*Market-to-book ratio*), for the previous year's cash flow (*Cash flow-to-assets ratio*), and for lagged firm size (ln(*Total assets*)).[10] We include year fixed effects (δ_t) to control for time varying factors common to all firms. Because equation (5) uses estimates of ρ, we calculate the standard errors conservatively by bootstrapping and clustering at the three-digit SIC level.

Table 6 presents the results. As predicted by theory, the estimate of β_ρ in column 1 is negative and significantly different from zero at the 1% level. This result is novel because the *inverse* relation between shock correlation and cash holdings is uniquely predicted by the combination of correlated permanent and short-term cash flow shocks. The estimate of β_ρ is also economically significant: A one sample standard deviation increase in $\hat{\rho}$ is associated with a decrease in the *Cash holdings* ratio of 4.7% relative to the sample mean. The coefficients of the control variables are similar to those found in the literature (see, e.g., Bates et al. (2009)). For example, the estimate of *Industry cash flow volatility* is positive and significant. However, the estimate of ρ contains variation in the firms' cash flow risk that is informative about cash policy over and above the information in the standard measure of industry cash flow volatility.

[10]The results are robust to the inclusion of other control variables such as capital expenditures, R&D, leverage, and dividends.

Next, we test the prediction that the firm's cash holdings increase with a higher permanent shock volatility using the regression model in (5) but replacing the estimated shock correlation with the estimated permanent shock volatility. Column 2 of Table 6 confirms this prediction, as the estimate of β_{σ_P} is positive and statistically and economically significant. A one sample standard deviation increase in $\hat{\sigma}_P$ is associated with an increase in cash holdings of 13% relative to its mean. In column 3, we study the relation between cash holdings and the volatility of short-term cash flow shocks, σ_A. As predicted, the estimated short-term cash flow volatility, $\hat{\sigma}_A$, also has a positive and significant effect on cash holdings. The economic significance of this volatility is smaller than that of the permanent shock volatility: A one sample standard deviation increase in $\hat{\sigma}_A$ is associated with a 6.4% increase in cash holdings relative to its mean. Finally, in column 4 we test the prediction that cash holdings increase with the growth rate of productivity (μ). This is exactly what we find: The estimate of β_μ is positive and statistically significant. A one standard deviation increase in $\hat{\mu}$ is associated with a 6.2% increase in cash holdings.

Insert Table 6 Here

In column 5, we include all the cash flow risk parameters, $\hat{\rho}$, $\hat{\sigma}_P$, and $\hat{\sigma}_A$ in the same regression model, while in column 6 we jointly include $\hat{\rho}$ and $\hat{\mu}$. In both columns, we also include two-digit SIC industry fixed effects.[11] In column 5, the coefficient of β_ρ is negative and statistically significant, while the coefficients of β_{σ_P} is positive and significant. That is, the within two-digit SIC industry variation of $\hat{\rho}$ and $\hat{\sigma}_P$ explain well cash policy heterogeneity *within* the industry. Moreover, the economic significance of the estimates is largely unaffected by the inclusion of industry fixed effects, a result that contrasts with the large drop from 14.8% to 3.5% in the estimated economic significance of *Industry cash flow volatility*. Importantly, the results indicate that our estimates of the correlations between permanent and short-term cash flow shocks explain substantial differences in cash holdings within industries that the standard measure of cash flow volatility, computed at the industry level,

[11] Controlling for the *Market-to-Book ratio*, which is typical for cash regressions in the empirical corporate finance literature, is problematic in this context because, in theory, Tobin's Q is a function of the cash flow shock volatilities, their correlation, and the growth rate. To address this problem, the specifications in columns 5 and 6 control instead for the residuals of the regression of the *Market-to-Book ratio* on $\hat{\rho}$, $\hat{\sigma}_P$, $\hat{\sigma}_A$, and $\hat{\mu}$.

cannot explain as well. The estimate of β_{σ_A} is also positive although no longer statistically significant. A potential reason for this result could be the relatively high correlation between $\hat{\sigma}_P$ and $\hat{\sigma}_A$ (the correlation coefficient is 0.421). In column 6, the coefficient of β_ρ is negative and statistically significant while the coefficient of β_μ is positive and statistically significant, confirming the results without industry fixed effects.

In untabulated results, we show that the empirical relation between cash holdings and $\hat{\rho}$, $\hat{\sigma}_P$, $\hat{\sigma}_A$, and $\hat{\mu}$ holds not only within each two-digit SIC industry but also cross-sectionally between all firms in our sample (using a Fama–MacBeth and a between estimator).

4.3 Equity issues

If raising cash by issuing new equity is an alternative to retaining earnings and holding cash, we would expect firms to issue more equity when the growth rate of productivity is higher and when short-term and permanent cash flow shock volatilities are higher. Similarly, we expect firms to issue less equity the higher the correlation between short-term and permanent shocks. These predictions mirror those for cash holdings. We therefore estimate the same regression model as for cash holdings, but replace the dependent variable with the total annual amount of equity issued by a firm:

$$\textit{Equity issuance}_{i,t} = \beta_0 + \delta_t + \beta_\rho \times \hat{\rho} + \boldsymbol{\beta}_{\text{Controls}} \times \textbf{Controls}_{i,t-1} + u_{i,t}. \tag{6}$$

We define *Equity issuance* as in McLean (2011). This variable captures all share issuances that result in cash flow to the firm (including proceeds from SEOs, private placements, rights offerings, stock sales through direct purchase plans, preferred stock issues, the conversion of debt and preferred stock, and employee options, grants, and benefit plans) and is scaled by lagged total assets. The other variables are defined as in the previous regressions. Table 7 presents the results.

Insert Table 7 Here

The estimate of β_ρ is negative and significantly different from zero in columns 1, 5, and 6, as predicted by theory. A one sample standard deviation increase in $\hat{\rho}$ is associated with a decrease in *Equity issuance* of 6.8% relative to the sample mean (column 5). Moreover, columns 2, 3, and 5 show a positive and significant relation between equity issues and short-term and permanent cash flow shock volatility. A one sample standard deviation increase in $\hat{\sigma}_P$ or $\hat{\sigma}_A$ is associated with an increase in equity issuance of 13.0% and 7.7%, respectively (column 5). Finally, the coefficient of the growth rate of permanent shocks is positive, as predicted, if statistically significant only at the 10% level (column 6).

4.4 Credit lines

Firms can use credit lines as an alternative to cash reserves for liquidity management. Theory predicts that the required size of a credit line should be negatively related to the correlation between permanent and short-term cash flow shocks, but positively related to the volatility of either shocks and to the growth rate of productivity. To test these predictions, we estimate the following regression:

$$Credit\ lines_{i,t} = \beta_0 + \delta_t + \beta_\rho \times \hat{\rho} + \boldsymbol{\beta}_{\text{Controls}} \times \mathbf{Controls}_{i,t-1} + u_{i,t}, \tag{7}$$

which includes the same control variables as previous regressions. The dependent variable is defined as the amount of the outstanding credit lines for a firm in a given year, as a proportion of the firm's borrowing amount, i.e., total debt. Credit lines include revolvers, lines of credit (with maturities below and above one year), and 364-day facilities (see, e.g., Sufi (2009)). We divide credit lines by total debt to scale credit lines by a measure of the firms' debt capacity. Table 8 presents the results.

Insert Table 8 Here

Consistent with theory, the coefficient estimate of β_ρ is negative and significant in all columns of Table 8. For example, in column 5, a one sample standard deviation increase in $\hat{\rho}$ is associated with a decrease in *Credit lines* of 7.7% relative to the sample mean. This result confirms our earlier findings for cash and equity issues that firms with negatively correlated short-term and permanent cash flow shocks need higher financial flexibility compared to firms for which these shocks are positively correlated.

The coefficient estimates of $\hat{\sigma}_P$ are all positive and significant and the economic magnitudes are large. A one sample standard deviation increase in $\hat{\sigma}_P$ is associated with an increase in the credit line to debt ratio of 16.6% (column 5). The estimates of $\hat{\sigma}_A$ are also positive, but their statistical and economic significance is slightly weaker. Finally, the estimates of $\hat{\mu}$ are positive, and statistically and economically significant, in line with theory.

4.5 Cash savings

Theory predicts that the cash flow sensitivity of savings is increasing in ρ and the ratio of σ_P/σ_A. In the Décamps et al. (2017) model, this sensitivity is proportional to $\rho \times \frac{\sigma_P}{\sigma_A}$, so that the sign of the cash flow sensitivity of cash is equal to the sign of ρ and its absolute value should be higher the higher the ratio of σ_P/σ_A. This analysis can shed light on the unsettled debate in the literature about whether the cash flow sensitivity of cash is positive (Almeida, Campello, and Weisbach (2004)) or negative (Riddick and Whited (2009)).

To test these predictions, we estimate the cash flow sensitivity of savings over several subsamples formed according to our estimates of ρ and the ratio σ_P/σ_A. We first define partitions of our sample data according to the estimates of $\hat{\rho}$. Because of estimation error in $\hat{\rho}$, we do not choose zero as the exact switching threshold of the cash flow sensitivity. Instead, we use the partition defined by values below, within, or above the empirical 99% coverage interval around zero, i.e., $\hat{\rho}_i \leq -0.02$, $\hat{\rho}_i \in (-0.02, 0.02)$, and $\hat{\rho}_i \geq 0.02$. For robustness, we also perform the tests over the alternative partition $\hat{\rho}_i \leq -0.03$, $\hat{\rho}_i \in (-0.03, 0.03)$, and $\hat{\rho}_i \geq 0.03$.

Sensitivity sign differences would be most clearly detected among firms with higher absolute sensitivities. Hence, we estimate the cash flow sensitivities of cash over increasingly restrictive subsamples based on the distribution of the ratio σ_P/σ_A, with values above the median, the 60th percentile, or the 70th percentile. The results are robust to measurements in the top tercile or quartile.

For each of the resulting nine subsamples (combining three sets of values for $\hat{\rho}$ and for $\hat{\sigma}_P/\hat{\sigma}_P$), we estimate the cash savings regression in Almeida et al. (2004) and Riddick and Whited (2009):

$$Cash\ savings_{i,t} = \beta_0 + \delta_t + \beta_{CF} \times Cash\ flow\text{-}to\text{-}assets_{i,t} + \boldsymbol{\beta}_{\text{Controls}} \times \textbf{Controls}_{i,t-1} + u_{i,t} \quad (8)$$

in which *Cash savings* is the yearly change in the stock of cash divided by total assets. Control variables include the *Market-to-book ratio* and ln(*Total assets*). Table 9 presents the estimates of the propensity to save obtained from this regression. Panel A presents coefficients estimated by OLS, including two-digit SIC industry fixed effects, and Panel B coefficients estimated using the fourth-order linear cumulants estimator (LC4) following Erickson, Jiang, and Whited (2014).

Insert Table 9 Here

The prediction that the cash flow sensitivity of cash switches sign is most clearly observed in the subsample of relatively high $\hat{\sigma}_P/\hat{\sigma}_A$ ratios (i.e., the top 30%). In this subsample, the estimated sensitivity, β_{CF}, is negative for $\hat{\rho}_i \leq -0.02$ (and statistically significant using an LC4 estimator) but positive and statistically significant when $\hat{\rho}_i \geq 0.02$. When the estimated ρ is around zero ($\hat{\rho}_i \in (-0.02, 0.02)$), the estimated cash flow sensitivity of cash is either not distinguishable from zero (OLS), or negative but closer to 0 than when $\hat{\rho}_i \geq 0.02$ (LC4). The results are even more pronounced when we partition the data at the -0.03 and 0.03 thresholds of $\hat{\rho}$, confirming that the switching of sign in the propensity to save is not driven by observations when $\hat{\rho}$ may be close to zero. Remarkably, the estimated sensitivity, β_{CF}, exhibits the predicted sign switch for all but one of the 12 sets of test partitions (LC4 estimator, $\hat{\sigma}_P/\hat{\sigma}_A$ above the median).

4.6 Additional analysis: SIC4 and random firm grouping

We perform two additional exercises to assess the appropriateness of our method of grouping firms with similar cash flow dynamics. First, we re-estimate the cash flow model in equations (1)–(2) using all firms in each four-digit SIC industry classification and then test whether these parameter estimates can better explain cash holdings, equity issuance, and credit line usage. The grouping at the four-digit SIC level is less granular and, as shown in Table 3, produces more heterogeneous groups of firms. We find that the cash flow model parameters estimated at the four-digit SIC level have similar distributions to those estimated more precisely using sorting by historical cash flow growth rates within three-digit SIC codes. However, the former do not explain the same liquidity policies as well as the latter. Overall, the regression coefficients do have the predicted signs but their statistical and economic significance are markedly weaker.

Second, we conduct a placebo test and re-estimate the model parameters in (1)–(2) for each of ten thousand groups of ten firms selected at random and with replacement from our data set. Again, we test whether the cross-sectional variation in these estimates explains differences in cash holdings, equity issuance, and credit lines. In this case, the parameter estimates are not related to corporate policies in any clear and systematic way.[12] We conclude that our grouping method captures the essential similarities of the firms' cash flow dynamics, and the parameter estimates measure deep characteristics of firms' cash flow risk.

5 Cash flow shock correlation and industry structure

Our estimates of permanent and short-term shock volatilities explain cash savings, equity issuance, and credit line policies as predicted by existing theories of optimal liquidity management in the presence of either transitory (Bolton et al., 2011; Décamps et al., 2011; Gryglewicz, 2011; Bolton et al., 2013; Hugonnier et al., 2015) or permanent cash flow shocks (Décamps et al., 2017). But our results also verify the *unique predictions* when both types of shocks are present and *correlated* as in Décamps et al. (2017). According to our evidence, more correlated permanent and short-term shocks imply lower cash holdings, lower equity issuance and smaller credit lines. This correlation uncovers a new feature of cash flow data, unexplored so far, and has an economically important impact on policies as predicted by theory. We conclude our analysis by discussing potential explanations for the deeper meaning and the determinants of the correlation parameter.

[12]We implement two different ways to merge randomly formed groups to Compustat firms. In one implementation, we match each Compustat firm to the parameter estimates of only one randomly selected group the firm is in. In another, we randomly select 794 of the 10,000 groups such that no firm belongs to more than one group. Neither implementation can reproduce the results under our proposed grouping.

5.1 Shock correlation across industries

The sign and absolute magnitude of shock correlations may be a characteristic intrinsic to each industry, varying mostly across industries but not across firms within each industry. Under this interpretation, for example, negative values of ρ may be pervasive in high-growth industries as firms face temporarily disruptions while adopting productivity-boosting discoveries. But contrary to this narrative, Table 4 shows that ρ varies mostly within rather than between three-digit SIC industries.

Insert Table 10 Here

To explore further, Table 10 decomposes the between and within industry variation of ρ in more detail. The distributions of the estimates of ρ across the 17 Fama–French industries (FF17) are remarkably similar: All of the medians are negative, ranging between -0.12 and -0.07. Except for three FF17 industries (Textiles, Consumer durables, and Fabricated products), fewer than 25% of the firms in each industry have a positive ρ. However, at least 5% of firms in each FF17 industry have a positive ρ. Moreover, Figure 3 shows that, for most industries, groups of firms exist with significantly high positive estimates of ρ. In sum, the estimates of ρ, which have been estimated independently for small groups of firms within each industry, have very similar distributions across very different industries. This finding suggests that the shock correlation is not an industry-specific characteristic but rather a variable capturing how firms perform within their industry.

Insert Figure 3 Here

5.2 Shock correlation within industries

Next, we ask whether the estimated correlations are associated with firm characteristics within each industry. Table 11 reports the average Spearman's rank correlation coefficients between the estimates of ρ and basic industry outcome variables, corporate policies, and other selected measures of industry leadership. All ranks and correlations are computed annually by industry and reported as an average over all industry-years. For robustness, we compute the rank correlations for different industry definitions: Three- and four-digit SIC codes, and the 10, 12, 17, 30 and 38 Fama–French industries.

Table 11 shows that firms with the highest estimated shock correlations tend to have the largest market share, size and profitability within their industry. Higher estimates of ρ are also most frequent amongst the most levered yet most distant from default firms and those that invest more as a ratio of their total assets. The results are robust to all industry definitions. However, the strongest rank correlations are obtained when using the narrower distributions of industries, i.e., 17 or more Fama–French industries, or three- and four-digit SICs. High ρ firms are also the most innovative of their industry. They have the highest R&D to sales ratios and accumulate the most patents. Based on the market valuation index of Kogan et al. (2017), their patent portfolios are also the most valuable. Finally, the De Loecker, Eeckhout, and Unger (2018) estimate of the firm's markup is also higher for higher ρ firms in each industry.

Insert Table 11 Here

To summarize, the correlation between permanent and short-term shocks is not a characteristic intrinsic to each industry but rather a feature specific to small groups of firms and how they perform within their industry. High shock correlation appears to be associated with systematic high performance and innovation leadership. Moreover, in simulations of the Décamps et al. (2017) model, we have confirmed that firms with high ρ outlast otherwise similar firms with lower values of ρ, despite the fact that high ρ firms may also simultaneously experience negative productivity and short-term shocks. Again, high ρ firms have a natural hedge advantage in that low (high) short-term cash flows tend to occur only when productivity is declining (improving) and the firm has lower (higher) needs for cash.

5.3 Discussion

Using the above results, we close this section with a discussion of the deeper interpretation of the correlation between permanent and short-term shocks. Consider first how the parameter ρ is identified. If the shock correlation is zero then the cross-sectional average of cash flows of firms in the estimation group, $\overline{A}_t$, approaches the asset productivity P_t, as the number of firms grows. If instead the shock correlation is positive, cash flows co-move more strongly over time as even the firm-specific short-term shocks become synchronized within the group, so that overall cash flows depend more heavily on the common permanent shocks. As a result, the estimator of the model's parameters will infer high positive ρ values for groups in which *all* constituting firms' cash flows vary *simultaneously* and at a similar but highly volatile average growth rate.

To corroborate this reasoning, Table 11 shows that the firms with the highest estimates of ρ are also those with the highest correlation with their industry's (unweighted) average cash flow, i.e., they have the highest industry 'cash flow betas'. This moment condition also explains why ρ estimates are strongly correlated with indicators of leadership within an industry, since we estimate common ρ coefficients for small groups of firms based on their long-run growth rates.

Collecting these empirical results, a plausible interpretation of the meaning of the parameter ρ is that its sign captures the leader vs. follower status of any firm in an industry. Short-term and long-term cash flow shocks become positively correlated as leaders effectively profit from consumer preference shocks (Klemperer (1995)) or technological change (Spence (1981)) immediately, whereas followers experience temporary cash flow shortfalls while imperfectly imitating new products, seeing their demand for existing products cannibalized by the leaders' new ones, or facing costs of adopting new technologies (Hayashi (1982)). Essentially, the shock correlation in our cash flow model seems to identify in the reduced form the group of firms with a first-mover advantage (Lieberman and Montgomery (1988)).

6 Conclusion

Dynamic corporate finance models provide novel and somewhat unexpected predictions about the responses of optimal liquidity and financing policies to permanent and transitory cash flow shocks. We show that these shocks can be disentangled using cash flow data and a filter of our design, while estimating the cash flow model's parameters with a high level of granularity across Compustat firms. Our empirical analysis also shows that real-world firms actually implement policies that are in line with the models' predictions, in an economically meaningful way. The unique predictions of such model that the correlation between permanent and short-term shocks reduces the need for precautionary cash savings, credit lines, or large equity issuances, are verified. Moreover, this correlation seems to be a fundamental firm-specific feature associated to its leadership and performance in each industry.

Future research in corporate finance can explore the relation between our estimates of the cash flow model's parameters and other corporate policies or valuation. Our estimates can also be used to improve calibration or structural parameter estimation of dynamic corporate finance models.

Appendix

A. Summary of the Décamps et al. (2017) model

In this section, for the sake of completeness, we recall the cash flow model of Décamps et al. (2017) that features potentially correlated permanent and short-term cash flow shocks and is used to derive theoretical predictions about financing policies.

A.1 Cash flow model

In the continuous time model of Décamps et al. (2017), operating revenue is subject to permanent and transitory shocks. Asset productivity $P = (P_t)_{t\geq 0}$ is governed by the geometric Brownian motion:

$$dP_t = \mu P_t dt + \sigma_P P_t dW_t^P \tag{A1}$$

where μ and $\sigma_P > 0$ are constant and $W^P = (W_t^P)_{t\geq 0}$ is a standard Brownian motion. Therefore, asset productivity is non-stationary and features permanent shocks. In addition to these shocks, cash flows are subject to short-term shocks. For a given firm, the cash flows dA_t are proportional to P_t but uncertain and governed by:

$$dA_t = P_t dt + \sigma_A P_t dW_t^A \tag{A2}$$

where $\sigma_A > 0$ is constant and $W^A = (W_t^A)_{t\geq 0}$ is a standard Brownian motion. W^A and W^P can be correlated with correlation coefficient ρ, in that

$$\mathbb{E}[dW_t^P dW_t^A] = \rho dt, \text{ with } \rho \in (-1, 1). \tag{A3}$$

The specification for cash flow dynamics in equations (A1) and (A2) nests those in traditional dynamic corporate finance models. If $\sigma_A = 0$, we obtain the model with time-varying profitability applied extensively in dynamic capital structure models (see Goldstein, Ju, and Leland (2001), Hackbarth, Miao, and Morellec (2006), or Strebulaev (2007)) and real-options models (see Abel and Eberly (1994), Carlson, Fisher, and Giammarino (2006), or Morellec and Schürhoff (2011)). If $\mu = \sigma_P = 0$, we obtain the stationary framework of dynamic agency models (see DeMarzo and Sannikov (2006) or DeMarzo, Fishman, He, and Wang (2012)) and liquidity management models (see Décamps, Mariotti, Rochet, and Villeneuve (2011), Bolton, Chen, and Wang (2011), or Hugonnier, Malamud, and Morellec (2015)).

The transition equation (1) is a simple Euler discretization of (A1). The measurement equation (2) is related to an Euler discretization of (A2) and it is obtained by setting $A_{i,t}$ equal to the cash flow accumulated over year t, for firm i. In (2), P_{t-1} and not P_t enters the error term $P_{t-1}\,\sigma_A\,\varepsilon_{i,t}^A$ for it to have zero mean.

With the above specification, the firm's cash flow over the time interval $[t, t+dt]$ is given by

$$dA_t = P_t dt + \sigma_A P_t(\rho dW_t^P + \sqrt{1-\rho^2}dW_t^T) \tag{A4}$$

where $W^T = (W_t^T)_{t\geq 0}$ is a Brownian motion independent from W^P. This decomposition implies that short-term cash flow shocks dW_t^A consist of transitory shocks dW_t^T and permanent shocks dW_t^P.

A.2 Management's optimization problem

Short-term shocks expose the firm to potential losses that can be covered using cash reserves or new equity financing. Specifically, management is allowed to retain earnings inside the firm and we denote by M_t the firm's cash holdings at any time $t > 0$. Cash reserves earn a rate of return $r - \lambda$ inside the firm, where $\lambda > 0$ is a cost of holding liquidity. The firm can also raise additional funds from investors. External equity financing is costly with a fixed cost ϕP_t and a proportional cost θ. The dynamics of cash reserves are then given by:

$$dM_t = (r-\lambda)M_t dt + \left(dt + \sigma_A \rho dW_t^P + \sqrt{1-\rho^2}dW_t^T\right) P_t + \frac{dE_t}{\theta} - d\Phi_t - dL_t \tag{A5}$$

where E_t, Φ_t, and L_t are non-decreasing processes that represents the cumulative gross external financing, the cumulative fixed cost of financing, and the cumulative dividend paid to shareholders. Equation (A5) is an accounting identity that shows that cash reserves increase with the interest earned on cash holdings (first term on the right hand side), with the firm's earnings (second term), and with net external equity (third and fourth terms) and decrease with payouts (last term).

Management chooses the cash savings/payout and equity financing policies to maximize shareholder value. There are two state variables for the firm's optimization problem: Profitability P_t and the cash balance M_t. We can thus write this problem as

$$V(p,m) = \sup_{L,E} \mathbb{E}_{p,m}\left[\int_0^\infty e^{-rt}(dL_t - dE_t)\right] \tag{A6}$$

where p and m denote realizations of P and M at time $t = 0$.

B. Kalman filter and maximum likelihood estimation

This section provides a detailed exposition of the model estimation approach used in Section 2. We first describe the state space model and then derive the Kalman filter to compute the likelihood of cash flow data.

B.1 The state space form

The state space model in (1)–(2) consists of a transition equation and a measurement equation. The transition equation describes the discrete-time dynamics of the latent state process, which is the unobserved asset productivity P_t. The measurement equation describes the relation between the state process and the observed cash flows of firms that share the same asset productivity. To facilitate the exposition, we use a standard notation in state space models, and present the model as if missing observations were absent (Appendix B.3 discusses how we handle missing observations).

Let X_t denote the asset productivity in year t, i.e., we set $X_t = P_t$. The transition equation (1) can be rewritten as

$$X_t = \Phi_X X_{t-1} + \omega_t \tag{B1}$$

where $\Phi_X = (1+\mu)$, $\omega_t = \sigma_P X_{t-1}\varepsilon_t^P$ and $\varepsilon_t^P \sim \mathcal{N}(0,1)$. Thus, $\omega_t \sim \mathcal{N}(0, Q_t)$, where $Q_t = \sigma_P^2 X_{t-1}^2$, and the error term ε_t^P is the permanent shock to cash flows.

Let $Z_{i,t}$ denote the cash flows of firm i in year t, i.e., we set $Z_{i,t} = A_{i,t}$, and $Z_t = (Z_{1,t}, \ldots, Z_{N,t})'$ be the $N \times 1$ vector collecting the cash flows of the N firms that share the same asset productivity, where $'$ denotes transposition. The measurement equation in (2) can be written in vector form as

$$Z_t = H_Z X_t + u_t \tag{B2}$$

where the i-th element is $Z_{i,t} = X_t + u_{i,t}$, $u_{i,t} = \sigma_A X_{t-1}\varepsilon_{i,t}^A$, and $\varepsilon_{i,t}^A \sim \mathcal{N}(0,1)$ is the short-term shock to cash flows. In (B2), $H_Z = \mathbf{1}$, where $\mathbf{1} = (1, \ldots, 1)'$.

In classic applications of state space models, u_t is merely a measurement error of X_t, and it is assumed to be uncorrelated with X_t. In contrast, because permanent and short-term shocks in model (1)–(2) are correlated, u_t and X_t turn out to be correlated. Specifically, the correlation between $u_{i,t}$ and X_t is equal to ρ and enters the short-term shock $\varepsilon_{i,t}^A = \rho\varepsilon_t^P + \sqrt{1-\rho^2}\varepsilon_{i,t}^T$, where $\varepsilon_{i,t}^T \sim \mathcal{N}(0,1)$ is the transitory shock, uncorrelated with ε_t^P. Thus,

$$Cov[X_t, u_{i,t}|X_{t-1}] = \mathbb{E}[\omega_t\, u_{i,t}|X_{t-1}] = \mathbb{E}[\sigma_P X_{t-1}\varepsilon_t^P\, \sigma_A X_{t-1}\varepsilon_{i,t}^A|X_{t-1}] = \rho\sigma_P\sigma_A X_{t-1}^2.$$

Collecting the transitory shocks of the N firms in $\varepsilon_t^T = (\varepsilon_{1,t}^T, \ldots, \varepsilon_{N,t}^T)'$, the error term $u_t = \sigma_A X_{t-1}(\rho\varepsilon_t^P\mathbf{1} + \sqrt{1-\rho^2}\varepsilon_t^T) \sim \mathcal{N}(0, \Omega_t)$, where $\Omega_t = \sigma_A^2 X_{t-1}^2(\rho^2\mathbf{1}\mathbf{1}' + (1-\rho)^2 I_N)$, and I_N is the $N \times N$ identity matrix.

The correlation between u_t and X_t makes the standard Kalman filter biased and inconsistent. To overcome the issue posed by the correlation between u_t and X_t, we transform the measurement equation as follows

$$\begin{aligned} Z_t &= H_Z X_t + u_t + J(X_t - \Phi_X X_{t-1} - \omega_t) \\ &= (H_Z + J)X_t - J\Phi_X X_{t-1} + u_t - J\omega_t \\ &= H_Z^* X_t + \Phi_X^* X_{t-1} + u_t^* \end{aligned} \tag{B3}$$

where $H_Z^* = H_Z + J$, $\Phi_X^* = -J\Phi_X$, $u_t^* = u_t - J\omega_t$, and J is a $N \times 1$ vector that will be defined shortly. In the first equation, the third term on the right hand side is zero by definition of the transition equation (B1). This means that the transformed measurement equation (B3) is an exact alternative representation of the measurement equation (B2). Importantly, the vector J is defined such that the transformed measurement error u_t^* is uncorrelated with X_t

$$Cov[X_t, u_t^*|X_{t-1}] = \mathbb{E}[\omega_t\, u_t^*|X_{t-1}] = \mathbb{E}[\omega_t u_t|X_{t-1}] - J\mathbb{E}[\omega_t^2|X_{t-1}] = 0. \tag{B4}$$

Solving the last equation for J gives $J = \mathbb{E}[\omega_t u_t|X_{t-1}]/\mathbb{E}[\omega_t^2|X_{t-1}]$. In the state space model (B1)–(B2), J takes a simple form, that is $J = \rho\sigma_A/\sigma_P\mathbf{1}$.

Plugging J in u_t^* clarifies why u_t^* is uncorrelated with X_t in (B3)

$$u_t^* = u_t - J\omega_t = \sigma_A X_{t-1}\varepsilon_t^A - \rho\frac{\sigma_A}{\sigma_P}\mathbf{1}\sigma_P X_{t-1}\varepsilon_t^P = \sigma_A X_{t-1}(\varepsilon_t^A - \rho\mathbf{1}\varepsilon_t^P) = \sigma_A X_{t-1}\sqrt{1-\rho^2}\varepsilon_t^T$$

where ε_t^T is by definition uncorrelated with X_t, and we used $\varepsilon_t^A = (\varepsilon_{1,t}^A, \ldots, \varepsilon_{N,t}^A)'$. The error term u_t^* is by definition uncorrelated with X_{t-1} too.

The transformation of the measurement equation in (B3) can be applied to more general state space models to handle the correlation between state variables and measurement errors. For example, if X_t is a $k \times 1$ state variable, then $J = \mathbb{E}[u_t\,\omega_t^\top|X_{t-1}]\,\mathbb{E}[\omega_t\,\omega_t^\top|X_{t-1}]^{-1}$, which is an $N \times k$ matrix. Also, J could be time varying when the conditional expectations above are state dependent.

In the signal processing literature, Ma, Wang, and Chen (2010) suggest to transform the transition equation to account for the correlation between measurement and transition errors in state space models. We use a different approach and transform the measurement equation which results in a stable Kalman filter for the state space model in (1)–(2).

B.2 The generalized Kalman filter

Because the transformed measurement equation (B3) features X_{t-1} in the right hand side, it is necessary to re-derive the Kalman filter to filter out the latent state process.

Let $\hat{X}_{t|t-1} = \mathbb{E}_{t-1}[X_t]$ and $\hat{Z}_{t|t-1} = \mathbb{E}_{t-1}[Z_t]$ denote the expectation of X_t and Z_t, respectively, using information up to and including time $t-1$, and let $V_{t|t-1}$ and $F_{t|t-1}$ denote the corresponding (a priori) error variance and error covariance matrix. Furthermore, let $\hat{X}_t = \mathbb{E}_t[X_t]$ denote the expectation of X_t including information at time t, and let V_t denote the (a posteriori) error variance.

The Kalman filter consists of two steps, i.e., prediction and update. In the prediction step, $\hat{X}_{t|t-1}$ and $V_{t|t-1}$ are given by

$$\hat{X}_{t|t-1} = \Phi_X\hat{X}_{t-1} \tag{B5}$$
$$V_{t|t-1} = \Phi_X V_{t-1}\Phi_X + Q_t. \tag{B6}$$

and $\hat{Z}_{t|t-1}$ and $F_{t|t-1}$ are in turn given by

$$\hat{Z}_{t|t-1} = H_Z\hat{X}_{t|t-1} + \Phi_X^*\hat{X}_{t-1} \tag{B7}$$
$$F_{t|t-1} = H_Z^* V_{t|t-1} {H_Z^*}' + \Phi_X^* V_{t-1} {\Phi_X^*}' + \Omega_t^*. \tag{B8}$$

Because the transition equation (B1) is standard, $\hat{X}_{t|t-1}$ and $V_{t|t-1}$ take the usual forms in Kalman filtering. The transformed measurement equation (B3) changes $\hat{Z}_{t|t-1}$ and $F_{t|t-1}$ relative to standard Kalman filtering, with the additional terms in Φ_X^*.

In the update step, the estimate of the state vector X_t is refined based on the difference between the observed and predicted values of Z_t, with $\hat{X}_t$ and V_t given by

$$\hat{X}_t = \hat{X}_{t|t-1} + G_t'(Z_t - \hat{Z}_{t|t-1}) \tag{B9}$$

$$V_t = V_{t|t-1} - 2V_{t|t-1}H_Z^{*\prime}G_t + G_t'F_{t|t-1}G_t \tag{B10}$$

where G_t is an $N \times 1$ vector called Kalman gain, which is determined by minimizing V_t with respect to G_t. Solving the first order condition $\partial V_t/\partial G_t' = 0$ for G_t, gives $G_t' = V_{t|t-1}H_Z^{*\prime}F_{t|t-1}^{-1}$. This choice of G_t minimizes V_t because $\partial^2 V_t/(\partial G_t\,\partial G_t') = 2F_{t|t-1}$ is positive definite.

Model estimation is achieved by maximizing the log-likelihood of cash flows data of N firms over T periods with respect to the model parameters μ, σ_P, σ_A, and ρ. Specifically, for fixed model parameters the generalized Kalman filter (B5)–(B10) is run to compute the log-likelihood

$$\sum_{t=1}^{T} -\frac{1}{2}\left[N\log(2\pi) + \log|F_{t|t-1}| + (Z_t - \hat{Z}_{t|t-1})'F_{t|t-1}^{-1}(Z_t - \hat{Z}_{t|t-1})\right]. \tag{B11}$$

Model parameters are changed as to increase the value of the log-likelihood, which then requires to re-run the generalized Kalman filter, and re-compute the log-likelihood. The iterative procedure is repeated until convergence of the numerical likelihood search. On a common laptop computer, it takes less than one second to fit the model to a panel of 10 firm cash flows observed over 46 years.

B.3 Missing observations handled with Kalman filtering

A prominent feature of cash flow data are missing observations. In our panel, 56% of firm-year observations are missing relative to the full balanced panel. Although our Kalman filter is different from the standard one, missing values can be handled using the usual method in Kalman filtering; see Section 3 in Shumway and Stoffer (1982). For completeness we briefly recall the procedure.

Suppose that there are no missing observations in year t. Then, the measurement equation (B3) holds. That is, Z_t collects the cash flows of all the N firms in a year t. Suppose now that the cash flow data of some firms in year t are missing. The idea is to "select" the components of Z_t corresponding to firms with observed (not missing) cash flow data. This task is achieved by simply using a matrix S_t consisting of zeros and ones with dimension $M_t \times N$, where M_t is the number of firms with observed cash flow data. To illustrate, consider an extreme and unrealistic case in which only the cash flow of the first firm in Z_t is available in year t. In that case, $S_t = (1, 0, \ldots, 0)$ is a $1 \times N$ row vector, $M_t = 1$ and $S_t\,Z_t$ is the cash flow of that firm. If cash flows of all N firms are available in year t, then S_t is a

$N \times N$ identity matrix.

The procedure to compute the log-likelihood with missing observations is as follows. First, for each year t, construct the matrix S_t based on the position of observed cash flows in Z_t. Then, pre-multiply both sides of the measurement equation (B3) by S_t and use this measurement equation to run the generalized Kalman filter. Finally, compute the log-likelihood in (B11) replacing N by M_t.

The matrix S_t is time dependent and needs to be computed for each year t. This time dependence allows the procedure to accommodate missing observations in different positions of the cash flow panel as well as entry and exit of firms in the panel.

B.4 Identification of shock correlation

To illustrate the issue of the identification of shock correlation, we use a popular model in the corporate finance literature. The model features persistent and short-term productivity shocks at a firm level. Denote $y_{i,t}$ an observable productivity or cash flow process for a given firm i,

$$y_{i,t} = \varepsilon_{i,t} + \sigma_\nu \nu_{i,t} \tag{B12}$$

where $\varepsilon_{i,t}$ is an AR(1) process, namely $\varepsilon_{i,t} = \beta\varepsilon_{i,t-1} + \sigma_\eta \eta_{i,t}$ and $0 < \beta < 1$ and $\sigma_\eta > 0$. The process $\eta_{i,t} \sim i.i.d.\mathcal{N}(0, 1)$ models persistent (or often called long-term) shocks. The process $\nu_{i,t} \sim i.i.d.\mathcal{N}(0, 1)$ models short-term shocks and $\sigma_\nu > 0$. Both shocks are firm-specific. The usual assumption in the literature is that these shocks are uncorrelated. We consider instead the case in which these shocks are correlated, $corr[\eta_{i,t}, \nu_{i,t}] = \rho$. It is perhaps surprising that even observing an infinite time series of $y_{i,t}$, the correlation ρ (and other model parameters) cannot be identified. Below we formally prove this result.

A time series model is identified when the system of equations, matching population and model-based autocovariances, can be solved uniquely for the model parameters. The unknowns in this system are the model parameters. Population autocovariances are (asymptotically) known. Define the autocovariance function as $\gamma(h) = Cov[y_{i,t}, y_{i,t-h}]$, for $h = 0, 1, \ldots$, then

$$\gamma(0) = \frac{\sigma_\eta^2}{1-\beta^2} + \sigma_\nu^2 + 2\rho\sigma_\eta\sigma_\nu \tag{B13}$$

$$\gamma(h) = \beta^h \left[\frac{\sigma_\eta^2}{1-\beta^2} + \rho\sigma_\eta\sigma_\nu\right]. \tag{B14}$$

The parameter β can be easily identified from the decay of the autocovariance function $\gamma(h)$, say from the equation $\gamma(h_2)/\gamma(h_1) = \beta^{h_2-h_1}$ for $h_2 > h_1 \geq 1$, and it is therefore taken as known in the discussion below. Although there is an infinite number of equations (B14) for $h \geq 1$, effectively, (B13)–(B14) is a system of two equations in three unknowns, $\rho, \sigma_\eta, \sigma_\nu$, and the model in (B12) is not identified.

To see the lack of identification of the shock correlation, suppose for simplicity that $\rho > 0$.

Solving (B13) for σ_ν (which then admits only one real and positive solution) and plugging this solution in (B14) gives

$$\gamma(h) = \beta^h \left[\frac{\sigma_\eta^2}{1-\beta^2} + \rho\sigma_\eta \left(\sqrt{\rho^2\sigma_\eta^2 + \left(\gamma(0) - \frac{\sigma_\eta^2}{1-\beta^2}\right)} - \rho\sigma_\eta \right) \right]$$

which is effectively one equation in two unknowns, ρ and σ_η. Therefore, ρ and σ_η are not identified.

Consider now the model

$$y_{i,t} = \varepsilon_t + \sigma_\nu \nu_{i,t} \tag{B15}$$

in which the persistent shock $\varepsilon_t = \beta\varepsilon_{t-1} + \sigma_\eta\eta_t$ is not firm-specific but common across firms, and is still correlated with the short-term shock, $corr[\eta_t, \nu_{i,t}] = \rho$. The persistent shock ε_t plays the role of a "systemic factor" for the firms' productivity. The short-term shock can be decomposed as $\nu_{i,t} = \rho\varepsilon_t + \sqrt{1-\rho^2}\nu_{i,t}^T$, where $\nu_{i,t}^T$ is the firm-specific transitory shock, uncorrelated with ε_t. The cross-sectional mean, $\overline{y}_t$, is such that

$$\overline{y}_t = \frac{1}{N}\sum_{i=1}^N y_{i,t} = \frac{1}{N}\sum_{i=1}^N (\varepsilon_t + \sigma_\nu\nu_{i,t}) = \frac{1}{N}\sum_{i=1}^N (\varepsilon_t + \sigma_\nu(\rho\varepsilon_t + \sqrt{1-\rho^2}\nu_{i,t}^T))$$

and therefore when $N \to \infty$, $\overline{y}_t = \varepsilon_t(1+\rho\sigma_\nu)$. The last equation indicates that if, for example, $\rho > 0$, then firms' productivity load more on the systemic factor ε_t relative to the case when $\rho = 0$. Importantly, the autocovariances of $\overline{y}_t$ provide additional moment conditions to identify the model in (B15). In essence, additional information from the cross-section of firms allows to identify the model and in particular the shock correlation. Denote $\overline{\gamma}(0) = V[\overline{y}_t]$, then

$$\overline{\gamma}(0) = \frac{\sigma_\eta^2}{1-\beta^2}(1+\rho\sigma_\nu)^2. \tag{B16}$$

The moment conditions (B13), (B14) and (B16) provide a system of three equations in three unknowns, $\rho, \sigma_\eta, \sigma_\nu$, to identify the model in (B15). This system can be solved as follows. Solving (B14) with respect to $\rho\sigma_\nu$ and plugging this solution in (B16) gives a quadratic equation in which σ_η is the only unknown. Ensuring that only one real and positive solution exists, identifies σ_η. The difference between (B13) and (B14) gives

$$\gamma(0) - \frac{\gamma(h)}{\beta^h} = \sigma_\nu^2 + \rho\sigma_\eta\sigma_\nu. \tag{B17}$$

Matching the expression of $\rho\sigma_\eta\sigma_\nu$ from (B17) and from (B14) gives a linear equation in which σ_ν^2 is the only unknown, identifying this parameter. Having identified σ_η and σ_ν, (B13) can be used to identify ρ.

In sum, a model in which persistent and short-term shocks are both firm-specific is not identified. Instead, assuming that persistent shocks are common across firms allows to identify the shock correlation, because these common shocks would behave like a systemic factor for firms' productivity.

C. Monte Carlo analysis of estimation accuracy

C.1 Standard vs. generalized Kalman filter

To check the accuracy of our estimation method, we conduct a Monte Carlo simulation. In the cash flow model given by equations (1) and (2), we set the parameters ρ, σ_A, σ_P, and μ to their respective average estimated values as in Table 4. We then use the model to simulate 10,000 panels of cash flows. As in our empirical analysis with Compustat data, each simulated panel consists of the cash flows of 10 firms over 46 years. For each simulated panel we estimate the model in (1)–(2) using maximum likelihood with our generalized Kalman filter, as described in Appendix B.2. As a benchmark method, we also estimate the cash flow model using maximum likelihood with a standard Kalman filter.

The standard filter presumes that the correlation between permanent and short-term cash flow shocks is zero, and thus delivers no estimate of ρ. As a measure of estimation accuracy, for each estimated parameter $\hat{\theta} = \{\hat{\rho}, \hat{\sigma}_A, \hat{\sigma}_P, \hat{\mu}\}$, we compute the mean square error (MSE), i.e., $\sum_{j=1}^{10000}(\hat{\theta}_j - \theta_0)^2/10000$, where $\hat{\theta}_j$ is the estimate of the parameter θ based on the j-th simulated panel of cash flows and θ_0 is the true parameter value. To compare MSE's across parameters, we report the relative MSE, i.e., the MSE divided by the absolute value of θ_0.

Underscoring the accuracy of our estimation method, the MSE's of the maximum likelihood estimates of the parameters μ, σ_P and σ_A with the generalized Kalman filter are, respectively, 0.360, 0.085 and 0.106. The MSE's of the maximum likelihood estimation with the standard Kalman filter are an order of magnitude larger than those with the generalized Kalman filter. The ratios between the two MSE's are 2.2, 1.7 and 2.1, respectively. Hence, our method is uniformly more accurate than maximum likelihood with a standard Kalman filter, often by a large extent. Finally, the MSE of ρ based on maximum likelihood with the generalized Kalman filter is 0.034, which is even smaller than the MSE's of the other parameters. As mentioned above, maximum likelihood with standard Kalman filter provides no estimate of ρ.

In sum, the Monte Carlo simulation above shows that maximum likelihood with the generalized Kalman filter delivers accurate estimates of the cash flow model in equations (1) and (2) while outperforming the maximum likelihood estimator with a standard Kalman filter. The latter method is not suited to handle the correlation between permanent and short-term shocks in (1)–(2).

C.2 Generalized Kalman vs. Hodrick-Prescott filters

Using the model (1)–(2), we simulate a panel of cash flows and then recover the asset productivity using the Kalman filer and the HP filter. We focus here on the HP filter because Byun et al. (2018) conduct an extensive Monte Carlo simulation and show that the HP filter is better suited than BN and Baxter–King filters to decompose firm cash flows into persistent and transitory shocks. Given a time series of observations, $y_t, t = 1, \ldots, T$, the HP filter is an additive decomposition $y_t = \tau_t + c_t$, where τ_t is identified as a trend component and c_t as a cyclical component. The trend component $\tau_t, t = 1, \ldots, T$, is obtained as the minimization of $\sum_{t=1}^{T}(y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1}((\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}))^2$, where the parameter $\lambda > 0$ controls the smoothness of the estimated trend component. The larger the value of λ, the smoother the resulting trend will be. The HP filter can be applied to a single time series to separate trend and cyclical components. Applying the HP filter to individual firm's cash flows would not recover the common (across firms) latent asset productivity, and would provide very inaccurate results. We apply the HP filter to the simulated group average of cash flows. Figure C1 shows the time series trajectory of the latent asset productivity (known in simulation), the Kalman-filtered asset productivity, and the HP-filtered trend component of the group average of cash flows.

Insert Figure C1 Here

The graph in Figure C1 indicates that the HP filter produces a "too smooth" trajectory of the asset productivity. This finding is to be expected because the HP filter is equivalent to a cubic spline smoother. Taking the HP-filtered trend component as latent asset productivity would lead to significant underestimation of the volatility of permanent shocks. In Figure C1, the volatility (i.e., sample standard deviation) of the true realized permanent shocks is 24%, but the volatility of HP-filtered permanent shocks is only 11%, which is more than a 50% underestimation of the realized volatility. Furthermore, because the HP filter underestimates substantially the volatility of permanent shocks, it largely overestimates the volatility of transitory shocks. In contrast, the Kalman-filtered asset productivity tracks closely the true asset productivity. The volatility of permanent shocks based on the Kalman filter is 29%, which is substantially closer to the true volatility than the HP-filtered volatility. The accuracy of the permanent shock volatility leads to a proper assessment of volatilities of short-term shocks. Firm-specific volatilities of short-term shocks range from 10% to 16%, Kalman-filtered volatilities are of similar magnitude, while HP-filtered volatilities are all above 60%.

References

Abel, A., and J. Eberly. 1994. A unified model of investment under uncertainty. *American Economic Review* 84:1369–1384.

Almeida, H., M. Campello, and M. Weisbach. 2004. The cash flow sensitivity of cash. *Journal of Finance* 59:1777–1804.

Bansal, R., R. Dittmar, and D. Kiku. 2008. Cointegration and consumption risks in asset returns. *Review of Financial Studies* 22:1343–1375.

Bates, T., K. Kahle, and R. Stulz. 2009. Why do U.S. firms hold so much more cash than they used to? *Journal of Finance* 64:1985–2021.

Bharath, S., and T. Shumway. 2008. Forecasting default with the Merton distance-to-default model. *Review of Financial Studies* 21:1339–1369.

Blundell, R., L. Pistaferri, and I. Preston. 2008. Consumption inequality and partial insurance. *American Economic Review* 98:1887–1921.

Boehmer, E., and J. Wu. 2013. Short selling and the price discovery process. *Review of Financial Studies* 26:287–322.

Bolton, P., H. Chen, and N. Wang. 2011. A unified theory of Tobin's Q, corporate investment, financing, and risk management. *Journal of Finance* 66:1545–1578.

Bolton, P., H. Chen, and N. Wang. 2013. Market timing, investment, and risk management. *Journal of Financial Economics* 109:40–62.

Brennan, M., and A. Subrahmanyam. 1996. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics* 41:441–464.

Byun, S., V. Polkovnichenko, and M. Rebello. 2018. Composition of cash flow shocks and debt financing. Working paper, University of Texas.

Carlson, M., A. Fisher, and R. Giammarino. 2006. Corporate investment and asset price dynamics: Implications for SEO event-studies and long-run performance. *Journal of Finance* 61:1009–1034.

Chang, X., S. Dasgupta, G. Wong, and J. Yao. 2014. Cash-flow sensitivities and the allocation of internal cash flow. *Review of Financial Studies* 27:3628–3657.

Cocco, J., F. Gomes, and P. Maenhout. 2005. Consumption and portfolio choice over the life cycle. *Review of Financial Studies* 18:491–533.

Cochrane, J. 1994. Permanent and transitory components of GNP and stock prices. *Quarterly Journal of Economics* 109:241–265.

Cohen, R., P. Gompers, and T. Vuolteenaho. 2002. Who underreacts to cash-flow news? Evidence from trading between individuals and institutions. *Journal of Financial Economics* 66:409–462.

De Loecker, J., J. Eeckhout, and G. Unger. 2018. The rise of market power and the macroeconomic implications. NBER working paper.

Décamps, J.-P., S. Gryglewicz, E. Morellec, and S. Villeneuve. 2017. Corporate policies with permanent and transitory shocks. *Review of Financial Studies* 30:162–210.

Décamps, J.-P., T. Mariotti, J.-C. Rochet, and S. Villeneuve. 2011. Free cash flow, issuance costs, and stock prices. *Journal of Finance* 66:1501–1544.

DeMarzo, P., M. Fishman, Z. He, and N. Wang. 2012. Dynamic agency and the q theory of investment. *Journal of Finance* 67:2295–2340.

DeMarzo, P., and Y. Sannikov. 2006. Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance* 61:2681–2724.

Erickson, T., C. Jiang, and T. Whited. 2014. Minimum distance estimation of the error-in-variables model using linear cumulant equations. *The Journal of Econometrics* 183:211–221.

Fama, E. F., and K. R. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43:153–193.

Garleanu, N., L. Kogan, and S. Panageas. 2012a. Displacement risk and asset returns. *Journal of Financial Economics* 105:491–510.

Garleanu, N., S. Panageas, and J. Yu. 2012b. Technological growth and asset pricing. *Journal of Finance* 67:1265–1292.

Glosten, L., and L. Harris. 1988. Estimating the components of the bid/ask spread. *Journal of Financial Economics* 21:123–142.

Goldstein, R., N. Ju, and H. Leland. 2001. An EBIT-based model of dynamic capital structure. *Journal of Business* 74:483–512.

Gorbenko, A. S., and I. A. Strebulaev. 2010. Temporary versus permanent shocks: Explaining corporate financial policies. *Review of Financial Studies* 23:2591–2647.

Gottschalk, P., and R. Moffitt. 2009. The rising instability of U.S. earnings. *Journal of Economic Perspectives* 23:3–24.

Grenadier, S. R., and A. Malenko. 2010. A Bayesian approach to real options: The case of distinguishing between temporary and permanent shocks. *Journal of Finance* 65:1949–1986.

Gryglewicz, S. 2011. A theory of corporate financial decisions with liquidity and solvency concerns. *Journal of Financial Economics* 99:365–384.

Gryglewicz, S., S. Mayer, and E. Morellec. 2019. Agency conflicts and short- vs. long-termism in corporate policies. *Journal of Financial Economics* forthcoming.

Guay, W., and J. Harford. 2000. The cash-flow permanence and information content of dividend increases versus repurchases. *Journal of Financial Economics* 57:385–415.

Guiso, L., L. Pistaferri, and F. Schivardi. 2005. Insurance within the firm. *Journal of Political Economy* 113:1054–1087.

Hackbarth, D., J. Miao, and E. Morellec. 2006. Capital structure, credit risk, and macroeconomic conditions. *Journal of Financial Economics* 82:519–550.

Hackbarth, D., A. Rivera, and T.-Y. Wong. 2018. Optimal short-termism. Working paper, Boston University.

Hamilton, J. 1994. *Time series analysis.* Princeton, NJ: Princeton University Press.

Hayashi, F. 1982. Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica* 50:215–224.

Hoffmann, F., and S. Pfeil. 2010. Reward for luck in a dynamic agency model. *Review of Financial Studies* 23:3329–3345.

Hugonnier, J., S. Malamud, and E. Morellec. 2015. Capital supply uncertainty, cash holdings, and investment. *Review of Financial Studies* 28:391–445.

Kim, W., and M. Weisbach. 2008. Motivations for public equity offers: An international perspective. *Journal of Financial Economics* 87:281–307.

Klemperer, P. 1995. Competition when consumers have switching costs: An overview with applications to Industrial Organization, Macroeconomics, and International Trade. *Review Economic Studies* 62:515–539.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman. 2017. Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics* 132:665–712.

Kwiatkowski, D., P. Phillips, P. Schmidt, and Y. Shin. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54:159–178.

Lee, B.-S., and O. M. Rui. 2007. Time-series behavior of share repurchases and dividends. *Journal of Financial and Quantitative Analysis* 42:119–142.

Leland, H. E. 1994. Corporate debt value, bond covenants, and agency costs. *Journal of Finance* 49:1213–1252.

Leland, H. E. 1998. Agency costs, risk management, and capital structure. *Journal of Finance* 53:1213–1243.

Lieberman, M., and D. Montgomery. 1988. First-mover advantages. *Strategic Management Journal* 91:41–58.

Ma, L., H. Wang, and J. Chen. 2010. Analysis of Kalman filter with correlated noises under different dependence. *Journal of Information and Computational Science* 5:1147–1154.

Malamud, S., and F. Zucchi. 2019. Liquidity, innovation, and endogenous growth. *Journal of Financial Economics* 132:519–541.

McLean, D. 2011. Share issuance and cash savings. *Journal of Financial Economics* 99:693–715.

Meghir, C., and L. Pistaferri. 2004. Income variance dynamics and heterogeneity. *Econometrica* 72:1–32.

Morellec, E., and N. Schürhoff. 2011. Corporate investment and financing under asymmetric information. *Journal of Financial Economics* 99:262–288.

Riddick, L., and T. Whited. 2009. The corporate propensity to save. *Journal of Finance* 64:1729–1766.

Shumway, R., and D. Stoffer. 1982. An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis* 3:253–264.

Spence, M. 1981. The learning curve and competition. *Bell Journal of Economics* 12:49–70.

Strebulaev, I. A. 2007. Do tests of capital structure theory mean what they say? *Journal of Finance* 62:1747–1787.

Sufi, A. 2009. Bank lines of credit in corporate finance: An empirical analysis. *Review of Financial Studies* 22:1057–1088.

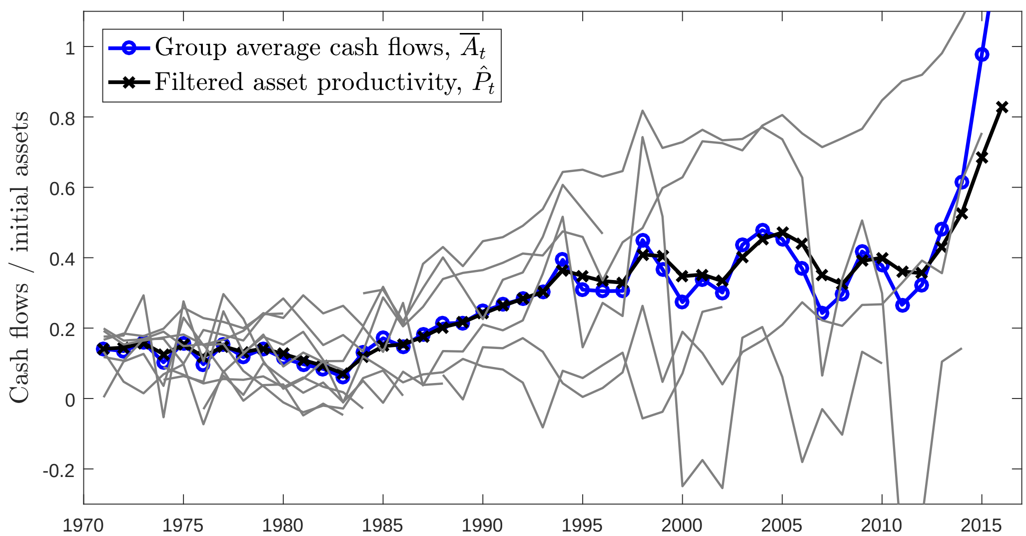

Figure 1: Cash flows for one group of firms. This figure shows the yearly firm cash flows scaled by the initial level of assets for a select group of ten firms in the 100 three-digit SIC code. The parameter estimates for the model in equations (1) to (2) are $\hat{\mu} = 0.08, \hat{\sigma}_P = 0.26, \hat{\sigma}_A = 0.10$, and $\hat{\rho} = -0.07$. The data covers the sample period 1971 to 2016.

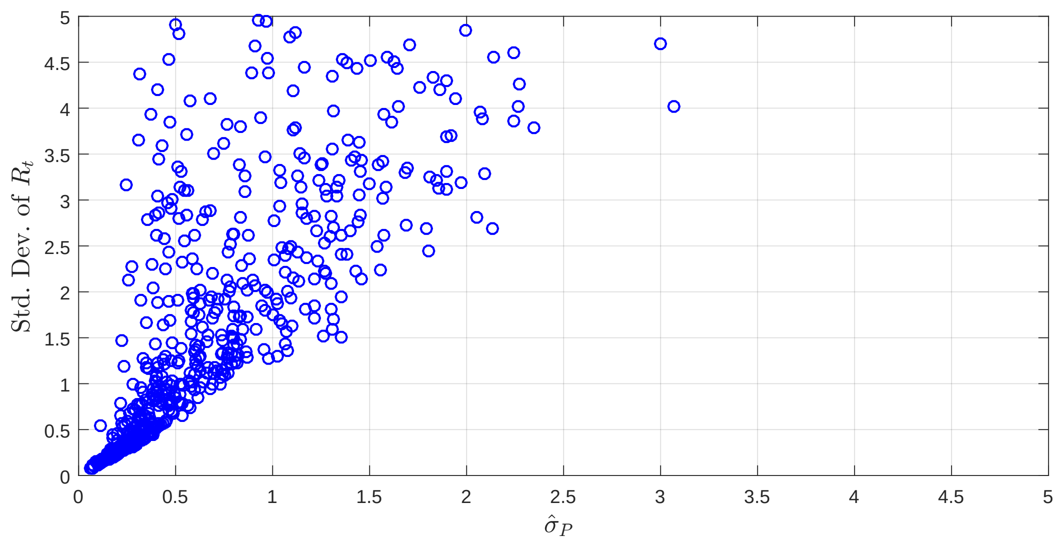

Figure 2: Scatter plot of permanent shock volatilities. The x-axis reports the model-based estimated volatilities of permanent shocks, σ_P in the model in (1)–(2), for each group of firms in our sample. The y-axis reports the observed time series standard deviations of R_t in (4) for $t = 1, \ldots, T$, i.e., the relative change of group-specific average cash flows. The sample data covers 9,232 firms, from 1971 to 2016, sorted in 794 groups.

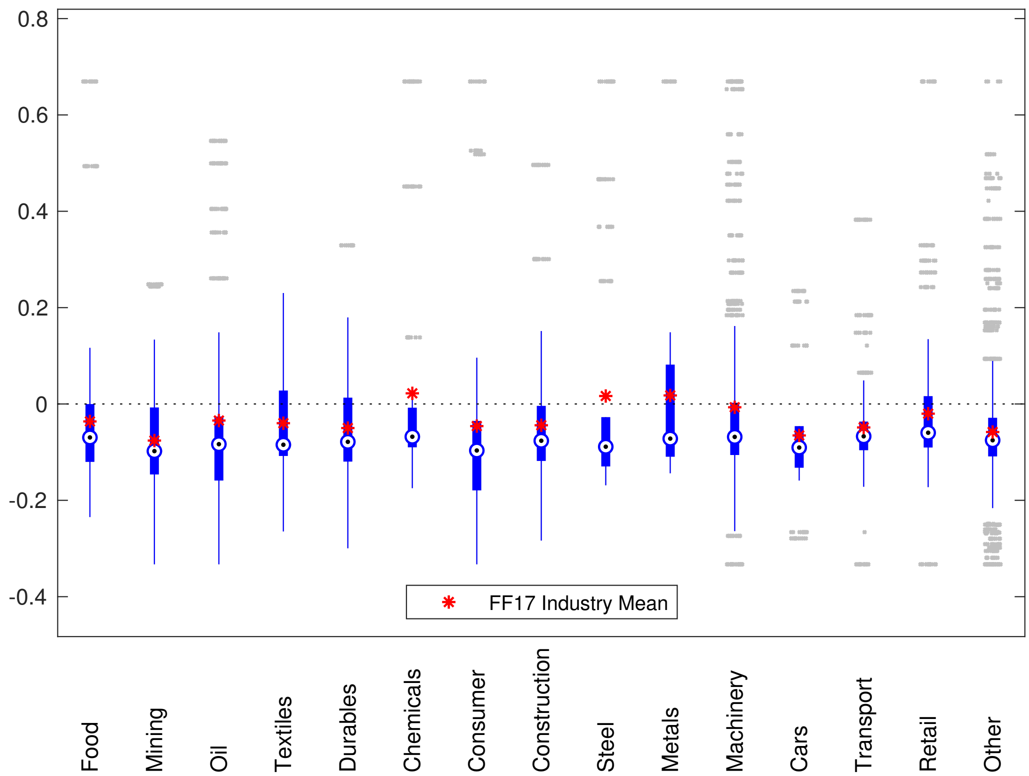

Figure 3: Estimates of the correlation between permanent and short-term cash flow shocks. This figure presents the box plots of the maximum likelihood estimates of the correlation coefficient, ρ, in the cash flow model of equations (1)–(2) for all firms within each of the 17 industries defined by Fama and French (1997). The sample data covers 9,232 firms, from 1971 to 2016.

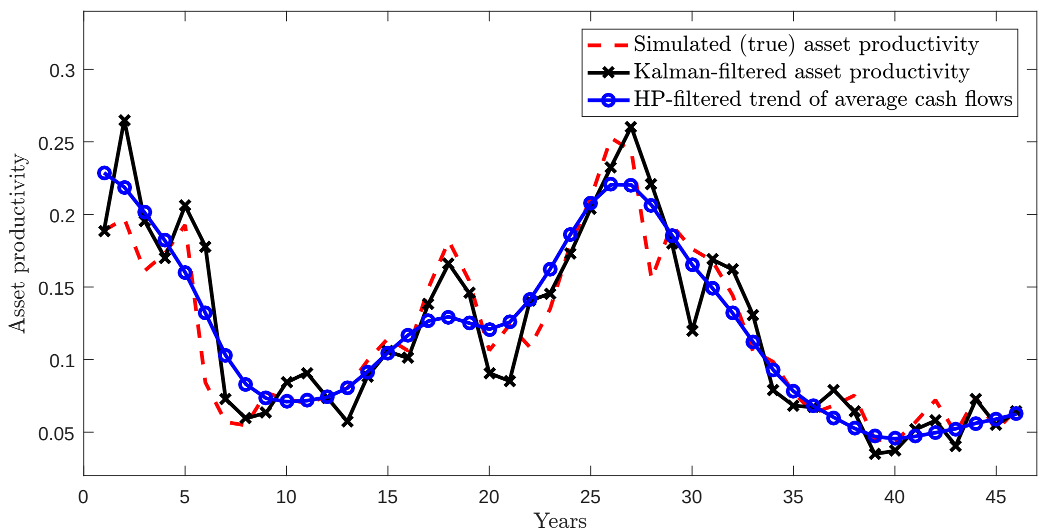

Figure C1: Generalized Kalman filter and Hodrick–Prescott filter applied to simulated cash flow data. Based on a panel of simulated cash flows from model in (1)–(2), the graph shows the time series trajectory of the true latent asset productivity, the Kalman-filtered asset productivity, and the trend component of the group average of cash flows from the HP filter. The Kalman-filtered asset productivity tracks closely the true asset productivity, while the HP filter, being a cubic spline smoother, provides a too smooth approximation of the asset productivity.

Table 1: Definitions and descriptive statistics of variables

This table presents the definitions and the descriptive statistics of the variables used in the analysis. The descriptive statistics are: Number of observations (N); mean; standard deviation, and the percentiles p5, p25, p50, p75, and p95. The sample covers the period 1971 to 2016.

Panel A: Variable definition

Variable name	Variable definition
Operating cash flow	EBITDA (oibdp) − change in working capital, defined as in Chang et al. (2014).
Cash flow-to-assets	Ratio of *Operating cash flow* to lagged book value of total assets (at)
Cash flow-to-initial assets	Ratio of *Operating cash flow* to the first observation of the book value of total assets (at)
Cash holdings	Cash and marketable securities (che) divided by the book value of total assets (at)
Cash savings	Ratio of the change in cash holdings (che) from year $t-1$ to t to the lagged book value of total assets (at)
Equity issuance	Ratio of the proceeds from sales or conversions of common and preferred stock (sstk) to the lagged book value of total assets (at)
Credit lines	Ratio of the total amount of credit lines outstanding (drawn and undrawn) to the total value of debt (dlc + dltt)
Market-to-book ratio	Book value of total assets (at) + market cap (csho*prcc_f) − book equity (ceq) divided by total assets
ln*(Total assets)*	Logarithm of the book value of total assets (at)
Industry cash flow volatility	Mean of the standard deviation of firms' cash flow-to-assets ratio, for all firms in the same two-digit SIC industry and year

Panel B: Descriptive statistics

	N	Mean	Stdev	p5	p25	p50	p75	p95
Cash flow-to-assets	176,474	0.078	0.253	−0.287	0.036	0.115	0.186	0.335
Cash flow-to-initial assets	186,008	0.725	2.447	−0.656	0.039	0.184	0.520	3.848
Cash holdings	186,542	0.140	0.179	0.002	0.022	0.069	0.185	0.544
Cash savings	176,905	0.025	0.192	−0.153	−0.023	0.000	0.032	0.236
Equity issuance	171,259	0.078	0.302	0.000	0.000	0.001	0.014	0.407
Credit lines	48,556	4.795	19.367	0.124	0.485	1.042	2.186	12.234
Market-to-book ratio	162,347	1.905	2.190	0.716	0.994	1.307	1.951	4.746
ln*(Total assets)*	186,656	5.071	2.359	1.300	3.403	5.013	6.686	9.147
Industry cash flow volatility	179,812	0.171	0.214	0.059	0.091	0.131	0.196	0.352

Table 2: Tests of non-stationarity of firms' cash flows

Table entries for ADF (columns 1 to 3) and for KPSS (columns 4 to 6) are the percentage of times that the ADF test is *not* rejected and the KPSS *is* rejected for three different confidence levels, 10%, 5%, 1%, respectively. The null hypothesis of the ADF test is that firm's cash flows have a unit root, i.e., they are non-stationary. The null hypothesis of the KPSS test is that the firm's cash flows are stationary. There are 9,232 firms in our cash flow panel. For any given firm there are maximum 46 yearly observations of cash flows from 1971 to 2016, and possibly less than 46 because of missing observations. The ADF and KPSS tests are run for each firm's cash flow time series, when the number of missing observations in its time series is below a given threshold (reported in Max NaN). Max NaN is the largest fraction of missing observations that are allowed for any given firm's cash flow time series to be considered for the ADF and KPSS tests. If Max NaN is zero, only firms' cash flows time series with no missing observations are considered for the ADF and KPSS tests. Critical values for the ADF and KPSS tests are not tabulated when Max NaN is larger than 2/3. # firms is the number of firms considered for the ADF and KPSS tests, which increases when more missing observations are allowed in firms' cash flow time series.

Panel A: *Cash flow-to-initial assets*

ADF			KPSS				
10%	5%	1%	10%	5%	1%	Max NaN	# firms
85.2	91.9	97.0	91.9	87.3	81.3	0	332
83.3	90.3	96.7	88.8	84.6	76.6	1/6	723
82.7	89.8	96.5	84.7	79.1	69.3	1/3	1,349
85.2	91.3	97.0	70.3	61.6	47.2	2/3	5,819

Panel B: *Cash flow-to-assets*

ADF			KPSS				
10%	5%	1%	10%	5%	1%	Max NaN	# firms
55.9	71.8	91.0	73.6	60.1	43.8	0	332
56.2	70.2	89.1	69.9	58.9	41.4	1/6	717
58.8	71.4	90.1	66.5	55.0	36.5	1/3	1,339
74.6	83.7	93.8	56.3	45.5	26.8	2/3	5,812

Table 3: Decomposition of standard deviation by industries or estimation groups

This table shows the decomposition of the total standard deviation of several firm-specific outcome, financing policy, product market and innovation variables into the between- and within-group standard deviations. Firms are grouped according to their 4-digit SIC code (SIC4), their 38-industry classifications in Fama and French (1997) (FF38), or allocated into groups of ten firms sorted by their average annual cash flow growth rate within each three-digit SIC code ('Groups'). The data is for all yearly observations of the 9,232 Compustat firms with at least 10 years of cash flow data between 1971 and 2016.

	Standard Deviation						
		Within-			Between-		
	Total	SIC4	FF38	Groups	SIC4	FF38	Groups
1. Outcome variables							
Annual sales-to-assets	1.53	1.38	1.45	1.38	0.61	0.45	0.69
Annual earnings-to-assets	0.28	0.26	0.27	0.26	0.06	0.05	0.10
Annual sales growth	0.46	0.46	0.46	0.46	0.07	0.04	0.08
2. Policy variables							
Cash holdings	0.18	0.16	0.17	0.16	0.06	0.05	0.08
Cash savings	0.19	0.19	0.19	0.19	0.02	0.01	0.03
Equity issuance	0.30	0.29	0.30	0.29	0.07	0.05	0.08
Credit lines	19.36	18.90	19.28	18.60	5.35	2.35	6.71
CAPEX-to-assets	0.10	0.10	0.10	0.09	0.04	0.03	0.04
Total debt-to-assets	12.11	12.08	12.11	12.03	0.91	0.24	1.45
3. Product market and innovation variables							
De Loecker et al. (2018) markups	0.94	0.83	0.89	0.83	0.33	0.30	0.42
R&D expense-to-assets	0.41	0.35	0.39	0.36	0.12	0.10	0.21
Number of patents	67.92	65.29	67.58	63.94	24.82	6.07	22.33
Market value of patents, as defined by Kogan et al. (2017)	965.78	958.51	963.30	938.12	103.35	94.64	205.55

Table 4: Summary of the parameter estimates of the cash flow model

This table summarises the maximum likelihood estimates of the cash flow model parameters, $\hat{\mu}, \hat{\sigma}_P, \hat{\sigma}_A, \hat{\rho}$,

$$\begin{aligned} P_t &= (1+\mu)\, P_{t-1} + \sigma_P\, P_{t-1}\, \varepsilon_t^P \\ A_{i,t} &= P_t + \sigma_A\, P_{t-1}\, \varepsilon_{i,t}^A \end{aligned}$$

where $\varepsilon_{i,t}^A = \rho\varepsilon_t^P + \sqrt{1-\rho^2}\varepsilon_{i,t}^T$, the correlation $\rho \in (-1, 1)$, P_t is the unobserved asset productivity, and $A_{i,t}$ are firm i cash flows in year t, for $i = 1, \ldots, N$ and $N = 10$. The permanent shock ε_t^P and transitory shock $\varepsilon_{i,t}^T$ are uncorrelated and distributed as $\mathcal{N}(0, 1)$. The model parameters are estimated for each of the 794 three-digit SIC-cash flow growth groups of firms. The descriptive statistics are: number of model parameter estimates (#est.); mean; (Total) standard deviation, decomposed into between- (sd_b), and within-three-digit SIC industry (sd_w) variation; and the percentiles p5, p25, p50, p75, and p95. The sample covers the period 1971 to 2016.

Panel A: Parameter estimates

	#est.	Mean	Standard deviation Total	sd_b	sd_w	p5	p25	p50	p75	p95
$\hat{\rho}$	794	−0.052	0.161	0.106	0.142	−0.249	−0.129	−0.081	−0.026	0.273
$\hat{\sigma}_P$	794	0.842	0.613	0.396	0.451	0.155	0.316	0.649	1.311	1.981
$\hat{\sigma}_A$	794	1.136	2.098	1.436	1.769	0.076	0.122	0.181	0.590	6.783
$\hat{\mu}$	794	0.164	0.119	0.081	0.096	0.026	0.066	0.128	0.237	0.398

Panel B: Absolute values of t-statistics of the parameter estimates

	#est.	Mean	Standard deviation	Proportion of p values < 0.05	< 0.01	p5	p25	p50	p75	p95
$\hat{\rho}$	794	27.950	54.913	0.580	0.533	0.260	0.912	3.356	19.328	209.933
$\hat{\sigma}_P$	794	85.392	177.146	0.796	0.697	1.215	2.778	7.948	56.147	697.622
$\hat{\sigma}_A$	794	6.182	7.403	0.849	0.765	0.958	2.238	3.912	6.298	32.097
$\hat{\mu}$	794	8.716	14.111	0.788	0.679	0.947	1.547	2.700	8.019	55.006

Panel C: Correlations between the parameter estimates

	$\hat{\rho}$	$\hat{\sigma}_P$	$\hat{\sigma}_A$	$\hat{\mu}$
$\hat{\rho}$	1.000			
$\hat{\sigma}_P$	-0.127^{***}	1.000		
$\hat{\sigma}_A$	-0.206^{***}	0.421^{***}	1.000	
$\hat{\mu}$	0.055	0.596^{***}	0.230^{***}	1.000

Table 5: Model-implied asset volatilities

This table presents a comparison of the distributions of empirical and model-implied asset volatilities. The volatilities are estimated for each of the 794 three-digit SIC-cash flow growth groups of firms. Model-implied asset volatilities are calculated using the model of Décamps et al. (2017) (see Appendix A) and the estimated cash flow parameters reported in Table 4. The remaining parameters are $r = 0.08$, $\lambda = 0.02$, $p = 1.06$, $\Phi = 0.002$, $\eta^P = \eta^T = 0.4$, and $\xi^T = \xi^P = 0.4$. Model-implied asset volatilities are winsorized at p5 and p95. The descriptive statistics are: Number of observations (N); mean; standard deviation, and the percentiles p5, p25, p50, p75, and p95.

	N	Mean	Stdev	p5	p25	p50	p75	p95
Empirical asset volatility	521	0.424	0.140	0.252	0.319	0.392	0.516	0.667
Model-implied asset volatility								
Baseline parameters	521	0.467	0.241	0.167	0.272	0.414	0.617	1.029
Baseline and $r = 0.05$	404	0.465	0.263	0.146	0.242	0.409	0.608	1.084
Baseline and $\Phi = 0.005$	517	0.444	0.226	0.163	0.252	0.403	0.582	0.963
Baseline and $\xi^T = \xi^P = 0.6$	506	0.417	0.183	0.173	0.278	0.395	0.518	0.770

Table 6: Cash holdings and cash flow shocks

This table presents estimates of regressions, where the dependent variable is the value of cash and marketable securities divided by the book value of assets (*Cash holdings*). The sample period is from 1971 to 2016. $\hat{\sigma}_P$, $\hat{\sigma}_A$, $\hat{\rho}$, and $\hat{\mu}$ are the estimated volatilities, correlations between permanent and short-term cash flow shocks, and the growth rate of asset productivity, common to all firms in the same SIC3–cash flow growth group. All specifications include year fixed effects, and specifications (5) and (6) include two-digit SIC industry fixed effects. Please refer to Table 1 for the definition of all the variables. Standard errors (in parentheses) are bootstrapped and clustered at the three-digit SIC industry level. Estimates followed by the symbols ***, **, or * are statistically significant at the 1%, 5%, or 10% levels, respectively.

Panel A: Coefficient estimates

Estimator	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) IFE	(6) IFE
$\hat{\rho}$	-0.042^{***}				-0.030^{**}	-0.046^{***}
	(0.016)				(0.014)	(0.017)
$\hat{\sigma}_P$		0.031^{***}			0.026^{***}	
		(0.005)			(0.005)	
$\hat{\sigma}_A$			0.005^{***}		0.001	
			(0.002)		(0.001)	
$\hat{\mu}$				0.077^{**}		0.076^{***}
				(0.037)		(0.018)
Industry cash flow volatility	0.144^{***}	0.126^{***}	0.139^{***}	0.141^{***}	0.034^{**}	0.035^{**}
	(0.048)	(0.048)	(0.048)	(0.049)	(0.017)	(0.017)
ln(*Total assets*)	-0.013^{***}	-0.011^{***}	-0.012^{***}	-0.012^{***}	-0.010^{***}	-0.011^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cash flow-to-assets	−0.062	−0.056	−0.061	−0.062	−0.053	−0.055
	(0.053)	(0.053)	(0.053)	(0.054)	(0.044)	(0.044)
Market-to-book ratio	0.014^{***}	0.014^{***}	0.014^{***}	0.014^{***}		
	(0.001)	(0.001)	(0.001)	(0.001)		
Market-to-book ratio residuals					0.012^{***}	0.012^{***}
					(0.002)	(0.002)
Observations	145,364	145,364	145,364	145,364	145,364	145,364
Adjusted R^2	0.159	0.167	0.161	0.160	0.212	0.210

Panel B: Economic significance calculated as: $\frac{\beta_x \times \text{Std. Dev.}(x)}{\text{mean}(\textit{Cash holdings})}$

	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\rho}$	-0.047^{***}				-0.034^{**}	-0.051^{***}
	(0.018)				(0.015)	(0.019)
$\hat{\sigma}_P$		0.129^{***}			0.107^{***}	
		(0.021)			(0.023)	
$\hat{\sigma}_A$			0.064^{***}		0.014	
			(0.024)		(0.016)	
$\hat{\mu}$				0.062^{***}		0.061^{***}
				(0.030)		(0.015)
Industry cash flow volatility	0.148^{***}	0.130^{***}	0.144^{***}	0.146^{***}	0.035^{**}	0.036^{**}
	(0.050)	(0.049)	(0.049)	(0.051)	(0.018)	(0.018)

Table 7: Equity issuance and cash flow shocks

This table presents estimates of regressions, where the dependent variable is *Equity Issuance*, measured as the ratio of the proceeds from sales or conversion of common and preferred stock to lagged book assets. The sample period is from 1971 to 2016. $\hat{\sigma}_P$, $\hat{\sigma}_A$, $\hat{\rho}$, and $\hat{\mu}$ are the estimated volatilities, correlations between permanent and short-term cash flow shocks, and the growth rate of asset productivity, common to all firms in the same SIC3–cash flow growth group. All specifications include year fixed effects, and specifications (5) and (6) include two-digit SIC industry fixed effects. Please refer to Table 1 for the definition of all the variables. Standard errors (in parentheses) are bootstrapped and clustered at the three-digit SIC industry level. Estimates followed by the symbols ***, **, or * are statistically significant at the 1%, 5%, or 10% levels, respectively.

Panel A: Coefficient estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	OLS	OLS	OLS	OLS	IFE	IFE
$\hat{\rho}$	−0.018**				−0.026***	−0.038***
	(0.008)				(0.008)	(0.008)
$\hat{\sigma}_P$		0.007**			0.013***	
		(0.003)			(0.003)	
$\hat{\sigma}_A$			0.002**		0.002***	
			(0.001)		(0.001)	
$\hat{\mu}$				−0.024		0.018*
				(0.023)		(0.010)
Industry cash flow volatility	0.019	0.015	0.017	0.020	−0.008	−0.008
	(0.015)	(0.015)	(0.015)	(0.016)	(0.012)	(0.012)
ln(*Total assets*)	−0.007***	−0.007***	−0.007***	−0.008***	−0.008***	−0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cash flow-to-assets	−0.235***	−0.234***	−0.235***	−0.236***	−0.224***	−0.224***
	(0.038)	(0.039)	(0.038)	(0.038)	(0.036)	(0.036)
Market-to-book ratio	0.036***	0.037***	0.036***	0.036***		
	(0.003)	(0.003)	(0.003)	(0.003)		
Market-to-book ratio residuals					0.036***	0.036***
					(0.002)	(0.002)
Observations	142,189	142,189	142,189	142,189	142,189	142,189
Adjusted R^2	0.250	0.250	0.250	0.250	0.258	0.258

Panel B: Economic significance calculated as: $\frac{\beta_x \times \text{Std. Dev.}(x)}{\text{mean}(\textit{Equity issuance})}$

	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\rho}$	−0.049**				−0.068***	−0.101***
	(0.022)				(0.022)	(0.022)
$\hat{\sigma}_P$		0.066**			0.130***	
		(0.030)			(0.032)	
$\hat{\sigma}_A$			0.062***		0.077***	
			(0.021)		(0.020)	
$\hat{\mu}$				−0.047		0.034*
				(0.044)		(0.019)

Table 8: Credit line size and cash flow shocks

This table presents estimates of regressions, where the dependent variable is *Credit lines*, measured as the ratio of the total amount of credit lines outstanding to total debt. The sample period is from 1982 to 2016. $\hat{\sigma}_P$, $\hat{\sigma}_A$, $\hat{\rho}$, and $\hat{\mu}$ are the estimated volatilities, correlations between permanent and short-term cash flow shocks, and the growth rate of asset productivity, common to all firms in the same SIC3–cash flow growth group. All specifications include year fixed effects, and specifications (5) and (6) include two-digit SIC industry fixed effects. Please refer to Table 1 for the definition of all the variables. Standard errors (in parentheses) are bootstrapped and clustered at the three-digit SIC industry level. Estimates followed by the symbols ***, **, or * are statistically significant at the 1%, 5%, or 10% levels, respectively.

Panel A: Coefficient estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	OLS	OLS	OLS	OLS	IFE	IFE
$\hat{\rho}$	−1.888**				−2.377**	−3.319***
	(0.931)				(0.942)	(0.960)
$\hat{\sigma}_P$		1.109***			1.494***	
		(0.359)			(0.410)	
$\hat{\sigma}_A$			0.264*		0.107	
			(0.139)		(0.143)	
$\hat{\mu}$				4.444**		5.707***
				(1.881)		(1.863)
Industry cash flow	0.948	0.419	0.690	0.709	−0.658	−0.661
volatility	(1.490)	(1.437)	(1.432)	(1.444)	(0.769)	(0.764)
ln(*Total assets*)	−1.532***	−1.446***	−1.491***	−1.476***	−1.327***	−1.360***
	(0.101)	(0.101)	(0.096)	(0.098)	(0.105)	(0.103)
Cash flow-to-assets	0.223	0.795	0.397	0.450	0.903	0.619
	(1.418)	(1.509)	(1.438)	(1.475)	(1.522)	(1.478)
Market-to-book	1.211***	1.185***	1.191***	1.215***		
ratio	(0.146)	(0.146)	(0.147)	(0.165)		
Market-to-book					1.042***	1.062***
ratio residuals					(0.182)	(0.186)
Observations	43,951	43,951	43,951	43,951	43,951	43,951
Adjusted R^2	0.027	0.028	0.028	0.028	0.037	0.036

Panel B: Economic significance calculated as: $\frac{\beta_x \times \text{Std. Dev.}(x)}{\text{mean}(\textit{Credit lines})}$

	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\rho}$	−0.061*				−0.077**	−0.108***
	(0.030)				(0.031)	(0.031)
$\hat{\sigma}_P$		0.123***			0.166***	
		(0.045)			(0.044)	
$\hat{\sigma}_A$			0.097*		0.039	
			(0.051)		(0.053)	
$\hat{\mu}$				0.097**		0.125***
				(0.041)		(0.041)

Table 9: Savings sensitivity of cash flow and cash flow shocks

This table presents estimates of the sensitivity of cash savings to cash flow, which are obtained from the slope coefficient of the regression of the yearly change in the stock of cash divided by total assets (*Cash savings*) on the firm's *Cash flow-to-assets*. Control variables include the lagged logarithm of *Total assets* and *Market-to-book ratio*. The sample period from 1971 to 2016. The data is sorted and classified into subsamples according to the ratio $\hat{\sigma}_P/\hat{\sigma}_A$, and $\hat{\rho}$, which are the ratio of the estimated volatilities of and correlations between permanent and short-term cash flow shocks, respectively, common to all firms in the same SIC3–cash flow growth group. The coefficients in Panel A are estimated by OLS, including two-digit SIC industry and year fixed effects. Standard errors are clustered at the firm level. The coefficients in Panel B are estimated using the fourth-order linear cumulants estimator (LC4) following Erickson et al. (2014) and standard errors are computed using the optimal GMM weighting matrix. Estimates followed by the symbols ***, **, or * are statistically significant at the 1%, 5%, or 10% levels, respectively. Please refer to Table 1 for the definition of all the variables.

	Panel A: OLS			Panel B: LC4		
	Subsamples by values of $\frac{\hat{\sigma}_P}{\hat{\sigma}_A}$:			Subsample of values of $\frac{\hat{\sigma}_P}{\hat{\sigma}_A}$:		
Subsamples by values of $\hat{\rho}$:	Above 50%	Above 60%	Above 70%	Above 50%	Above 60%	Above 70%
$\hat{\rho}_i \leq -0.02$	−0.007	−0.015	−0.016	−0.111***	−0.148***	−0.148***
$\hat{\rho}_i \in (-0.02, 0.02)$	0.007	0.006	−0.001	−0.148*	−0.150*	−0.164**
$\hat{\rho}_i \geq 0.02$	0.021	0.039*	0.046**	0.201***	0.233***	0.227***
$\hat{\rho}_i \leq -0.03$	−0.011	−0.019*	−0.021*	0.256***	−0.141***	−0.137***
$\hat{\rho}_i \in (-0.03, 0.03)$	0.009	0.009	0.017	−0.140*	−0.143*	−0.140
$\hat{\rho}_i \geq 0.03$	0.029	0.050**	0.049**	0.206***	0.236***	0.233***

Table 10: Estimates of the correlation between permanent and transitory cash flow shocks by industry

This table summarises the distribution of the maximum likelihood estimates of the correlation between permanent and transitory cash flow shocks, $\hat{\rho}$. This parameter is estimated, together with the other cash flow model parameters, for each of the 794 three-digit SIC-cash flow growth groups of firms, using firm-specific cash flow data from 1971 to 2016. The summaries show the number of firms, and the mean, standard deviations and percentiles p5, p25, p50, p75, and p95 of $\hat{\rho}$ for all firms in each industry of the 17-industry classification in Fama and French (1997).

Fama-French Industry	Number of firms	Mean	Standard Deviation	p5	p25	p50	p75	p95
Food	394	−0.040	0.173	−0.227	−0.120	−0.071	−0.040	0.494
Mining and Minerals	440	−0.085	0.151	−0.262	−0.169	−0.115	−0.039	0.244
Oil and Petroleum Products	694	−0.046	0.190	−0.225	−0.158	−0.096	−0.026	0.405
Textiles, Apparel and Footware	277	−0.051	0.114	−0.197	−0.114	−0.086	0.028	0.183
Consumer Durables	363	−0.053	0.140	−0.292	−0.126	−0.079	0.013	0.180
Chemicals	198	0.022	0.227	−0.175	−0.090	−0.068	−0.008	0.669
Drugs, Soap, Parfumes and Tobacco	334	−0.078	0.214	−0.305	−0.216	−0.097	−0.044	0.526
Construction and Construction Materials	497	−0.046	0.146	−0.234	−0.118	−0.078	−0.009	0.301
Steel Works	188	0.017	0.243	−0.169	−0.130	−0.089	−0.027	0.669
Fabricated Products	120	0.018	0.217	−0.144	−0.109	−0.072	0.082	0.669
Machinery and Business Equipment	1,398	−0.034	0.190	−0.244	−0.126	−0.082	−0.008	0.455
Automobiles	175	−0.084	0.145	−0.279	−0.159	−0.096	−0.068	0.234
Transportation	466	−0.057	0.115	−0.235	−0.104	−0.070	−0.041	0.184
Retail Stores	731	−0.034	0.147	−0.173	−0.107	−0.066	−0.015	0.273
Other	2,957	−0.068	0.125	−0.249	−0.129	−0.080	−0.039	0.161

Table 11: The correlation between permanent and transitory cash flow shocks and within-industry performance

This table shows the average within-industry-year Spearman's rank correlation coefficients between the firm's estimated correlation between permanent and transitory cash flow shocks, $\hat{\rho}$, and each of a set of firm-specific characteristics listed below. For each-industry year, firms are ranked according to the estimated ρ or each firm characteristic. Spearman correlations are averaged over all industry-years. Industries are defined using several classifications: three- and four-digit SIC codes, and the 10-, 12-, 17-, 30-, and 38-industry classifications in Fama and French (1997). The parameter ρ is estimated, together with the other cash flow model parameters, for each of the 794 three-digit SIC-cash flow growth groups of firms, using firm-specific cash flow data from 1971 to 2016. The Industry earnings beta is the slope of the regression of each firm's cash flow on the simple average cash flows for all firms in the industry.

	SIC classification		Fama and French (1997) classification				
	3-digit	4-digit	10	12	17	30	38
1. Outcome variables							
Annual sales	0.703	0.637	0.363	0.373	0.608	0.464	0.522
Total assets	0.705	0.641	0.359	0.370	0.604	0.465	0.521
Annual earnings-to-assets	0.753	0.698	0.363	0.369	0.636	0.488	0.542
2. Policy variables							
Total debt-to-assets	0.750	0.689	0.371	0.382	0.647	0.506	0.551
Bharath and Shumway (2008) distance to default	0.714	0.652	0.369	0.380	0.613	0.480	0.530
CAPEX-to-assets	0.768	0.713	0.389	0.399	0.651	0.526	0.566
3. Innovation and industry leadership variables							
De Loecker et al. (2018) markups	0.713	0.642	0.343	0.357	0.601	0.485	0.510
R&D expense-to-assets	0.606	0.511	0.281	0.245	0.568	0.385	0.401
Number of patents	0.615	0.523	0.289	0.243	0.581	0.386	0.403
Market value of patents, as defined by Kogan et al. (2017)	0.614	0.522	0.290	0.244	0.580	0.387	0.402
Industry earnings beta	0.643	0.595	0.282	0.293	0.554	0.381	0.457