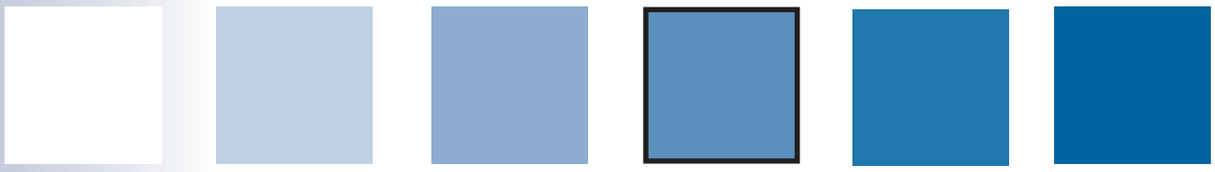
Federal Reserve Bank of St. Louis

REVIEW

July/August 2013 Volume 95, Number 4



REVIEW

Volume 95 • Number 4

Review is published six times per year by the Research Division of the Federal Reserve Bank of St. Louis and may be accessed through our website: **research.stlouisfed.org/publications/review**. All nonproprietary and nonconfidential data and programs for the articles written by Federal Reserve Bank of St. Louis staff and published in *Review* also are available to our readers on this website. These data and programs are also available through Inter-university Consortium for Political and Social Research (ICPSR) via their FTP site: **www.icpsr.umich.edu/pra/index.html**. Or contact the ICPSR at P.O. Box 1248, Ann Arbor, MI 48106-1248; 734-647-5000; **netmail@icpsr.umich.edu**.

Single-copy subscriptions (U.S. addresses only) are available free of charge: **https://research.stlouisfed.org/publications/review/subscribe/**.

General data can be obtained through FRED (Federal Reserve Economic Data), a database providing U.S. economic and financial data and regional data for the Eighth Federal Reserve District. You may access FRED through our website: **research.stlouisfed.org/fred**.

ISSN 0014-9187

The New Risk Management: The Good, the Bad, and the Ugly

Philip H. Dybvig, Pierre Jinghong Liang, and William J. Marshall

In a 1997 *Review* article, the authors described the good, the bad, and the ugly features of what they called the new risk management, which is the use of financial derivatives to hedge risk in firms. Since the article was first published, the "new" risk management has become commonplace and indeed played a big role in the financial crisis. As a result, the original article is more relevant today than when it was first published. This updated version of the article contains the same examples and critical analysis as in the original article but includes an updated description of the accounting rules and suggestions for designing a risk management policy. (JEL D81, G32, L21, M21, M41)

Federal Reserve Bank of St. Louis *Review*, July/August 2013, *95*(4), pp. 273-91.

At one time, risk management meant buying corporate insurance, implementing procedures to avoid lawsuits and accidents, and installing safety equipment. The new risk management uses financial markets to hedge different sources of risk within the firm. Companies can use trading in financial markets to hedge against the risk of changes in interest rates, input prices, or currency fluctuations. While hedging per se is not new, the scale and diversity of hedging are far greater than they used to be. When executed properly, the new risk management can be good and even essential for competition. Unfortunately, the new risk management can also be bad, wasting resources without reducing risk and perhaps even increasing it. The new risk management can be ugly, generating large losses such as those in widely publicized cases at Barings (in 1995), Metallgesellschaft (in 1993), Procter and Gamble (in 1994), and other firms. In these and many other firms, employees relatively far from the top of the hierarchy of control had the authority to take financial positions large enough to generate

Philip H. Dybvig is the Boatmen's Bancshares Professor of Banking and Finance at Washington University in St. Louis and a consultant to the Federal Reserve Bank of St. Louis. Pierre Jinghong Liang is an associate professor of accounting at Carnegie Mellon University. William J. Marshall is the president of NISA Investment Advisors, L.L.C. The authors are grateful for useful suggestions from Bill Gavin, Hong Liu, Myung-Yull Pang, Murray Weidenbaum, and Gaoqing Zhang. Philip Dybvig and Pierre Liang are grateful for support from the Southwestern University of Finance and Economics in Chengdu, Sichuan, China.

This article has been revised and updated since its original publication in *Review*: Dybvig, Philip H. and Marshall, William J. "The New Risk Management: The Good, the Bad, and the Ugly." Federal Reserve Bank of St. Louis *Review*, November/December 1997, *79*(6), pp. 9-22; http://research.stlouisfed.org/publications/review/97/11/9711pd.pdf.

losses that could bankrupt the firm. Thus, risk management policies should be put in place at the highest level of a firm and provide for monitoring and internal control of the amount of risk taken. The purpose of this article is to provide an introduction to the new risk management and some policy choices firms should be considering.

We start with a discussion of the option-pricing tools that make the new risk management possible and follow with a stylized example of how the new risk management ought to work. Then we consider implementation issues, including some general policy questions as well as some accounting issues.

TOOLS FOR THE NEW RISK MANAGEMENT

Starting with the famous work of Black and Scholes (1973; see the boxed insert on page 275), option-pricing theory has been very successful in pricing various financial claims. The Black-Scholes model was designed to price standard call and put options, and it has been extended to price all sorts of financial claims. The Black-Scholes model and its extensions form the theoretical foundation for the new risk management.

There were option-pricing models prior to the work of Black and Scholes, including some models with formulas similar to Black-Scholes. What makes the Black-Scholes model different is that it provides a *hedging strategy* that is an investment policy with an investment equal to the model's option price and a terminal value equal to the terminal value of the option. Knowing the hedging strategy is powerful since we can use this knowledge to make an arbitrage profit if market prices are out of line with the model. If the model price is lower than the price in the economy, we can sell the option, pocket the excess over the model price, and invest in the hedging strategy to cover the terminal value of the option we have sold. If the model price is higher than the price in the economy, we can follow the hedging strategy in reverse, taking a short position instead of a long position and lending instead of borrowing. In the model, the hedge replicates the option value perfectly; in practice, the hedge is not perfect, but it works remarkably well, which is why the Black-Scholes model and its progeny are widely used in business.

The introduction of these option-pricing models and the parallel development and maturation of liquid financial markets have made it easier and easier to hedge financial risks using options, futures, futures options, swaps, caps, collars, floors, and a variety of other financial instruments. (See the glossary of risk management terms in the appendix.)

OPTION PRICING AND RISK MANAGEMENT

Hedging an option is an example of risk management. Its purpose is to remove a transaction's risk and capture its pure economic profit. Fundamentally, this strategy is the same as insurance. For the insured, the insurance policy makes money in bad times (when an insurable event occurs) and loses money in good times (when no insurable event occurs but the premium is paid), which reduces risk by softening the impact of bad outcomes. The same is true of a hedging strategy; losing money on the hedge in good times and making money in bad times offsets the original cash flows, making the total cash flow less volatile. In either strategy, payment for the

The Black-Scholes Option-Pricing Model

The precursor of all modern option-pricing models was developed by Fischer Black and Myron Scholes (1973). The main result is an option-pricing formula based on simple and reasonable assumptions in a continuous-time model. The remarkable thing about the result is that it relies on the absence of arbitrage, and part of the proof is a formula that specifies a trading strategy in the underlying stock and the riskless bond that will replicate the payoff of the option at the end.* If the option is priced differently in the economy, buying or selling the option and following either the trading strategy or the reverse of the trading strategy will make money! Using the same sort of analysis, one can derive a trading strategy that will hedge the financial risk in a firm's cash flows.

Now we present the Black-Scholes formula for the price of a call option. Recall that a call option gives the owner the right (at the owner's option) but not the obligation to buy one share of the underlying stock at the strike (or exercise) price X specified in the option contract on or before the maturity date of the option. If the stock price is S and the price of the bond promising to pay the amount of the strike price at the maturity date of the option is B, the Black-Scholes price, C, of the call option is

$$C = S\,N(x1) - B\,N(x2),$$

where

$$x1 = \log(S/B)/s + s/2,$$

$$x2 = \log(S/B)/s - s/2, \text{ and}$$

s is the standard deviation (or square root of the variance) log of the stock price at maturity, given the stock price today, and the function $N()$ is the cumulative normal distribution function. If there is a constant, continuously compounded interest rate, r, and T is the time-to-maturity of the option, then B is the discounted exercise price

$$B = Xexp(-rT).$$

And, if the stock return has a variance, v, per unit time, we have that

$$s^2 = vT$$

is the variance log of the final stock price.

In the expression for C, the first term is the stock holding in the hedge strategy, and the second term is the bond holding (which is negative, which is a short sale or borrowing). The main assumptions of the model are absence of arbitrage, a constant riskless rate, continuous stock prices, and a constant variance of returns per unit of time for the underlying stock. The intuition is that we can replicate the risk of holding the option by holding just the right portfolio of riskless bonds and the underlying stock. For example, if at a point in time the option moves fifty cents for each one-dollar movement in the underlying stock price, then the replicating strategy would hold one share of stock for each two options we are replicating. To hedge the value of the option, we would short (borrow) a share of stock for each of two options. In that case, the stock's value change would neutralize the effect on our wealth of the option's price change. The hedge's holdings in the stock and bond will change over time and in response to stock price changes, since the sensitivity of the option value is different when the option is in the money than when it is out.

*For more discussion of why this makes sense, see Rubenstein and Leland (1981).

insurance can be "up-front" or "pay-as-you-go." For hedging, as for insurance, the arrangement of cash flows[1] accommodates the preference of the insured. There are important differences in taxation and regulation between hedging using insurance and hedging using financial markets, but those are beyond the scope of this article.

Using dynamic trading strategies to hedge financial options may seem significantly different from hedging price risk in a firm. However, the concept is exactly the same. A hedger is taking the other side of the risky investment in futures or whatever would be used to replicate the cash flows that are being hedged. Normally, these cash flows cannot be hedged precisely, but the hedge can still reduce risk significantly. For example, one policy is to hedge the expected cash flow conditional on the price of inputs that can be hedged in futures markets while leaving the remainder unhedged, which means that the remaining risk is borne by the stock and bond holders of the firm.

Before turning to the general policy issues in risk management, we consider a typical example.

RISK MANAGEMENT IN MANUFACTURING

Our example considers the hedging problem of a manufacturer that uses significant amounts of copper as an input. (With little change in the discussion, this input could be zinc, silver, oil, or wheat. With a slightly greater change, the "production" could be servicing of core deposits in a bank, and the analysis would provide the optimal hedging of interest rates.) We examine the optimal hedging of copper price movements in the cash flows before turning to a general discussion of policy and oversight.

In the example, the expected output is 1,000 units, which will sell for $100 per unit. The price has been committed to in advance because of long-term contracts, but the quantity may vary around this expectation because the contracts give customers the option to choose how much to buy within a range. Each unit will use an amount of copper that would cost $20 purchased forward (in a firm commitment to buy one year from now). If purchased in the spot market, the copper in the unit might cost $25 (with probability 1/4), $20 (with probability 1/2), or $15 (with probability 1/4).

One obvious (and common) approach to hedging in this context would be to forecast demand for copper and then hedge that amount, either by entering a fixed-price contract with the supplier or by buying that amount of copper futures—at a shorter maturity, if necessary, because one-year futures are not traded or have a very large spread. This might be a natural outcome if hedging were performed by buyers who were responsible for copper procurement and whose evaluations were based on the cost of a forecast quantity of copper. However, choosing a useful hedge of the entire cash flow is more subtle than that.

Table 1 contains an elaboration of the example. When the economy is doing well, copper prices are high (since this firm and other manufacturers are demanding more copper) and so is demand for the firm's output. Panel A of the table shows the cash flows in the absence of any special risk management to hedge copper price risk. Panel B shows the result of hedging by buying forward the expected quantity. Ironically, this naive approach to hedging increases risk exposure, since the firm is already more than hedged by increased sales when the industry is doing

Table 1

A Manufacturer's Copper Price Hedge

Each panel of this table shows the cash flows one year from now for the simple example of a manufacturer that is facing copper price risk. In the example, copper prices are higher when demand for output is higher. Each panel of the table illustrates a different hedging strategy and profit (= cash flow) for three copper price scenarios. The example abstracts from taxes and sources of risk that are not related to the price of copper. In each case, the expected profit is $1,500. The point of hedging is reducing uncertainty, not increasing average cash flow (except indirectly, because it allows more freedom in choosing projects).

Panel A: Unhedged Cash Flows

Probability	Copper price ($)	Units sold	Output price ($)	Total sales ($)	Copper expense ($)	Other expenses ($)	Profit (loss) ($)
¼	25	1,200	100	120,000	30,000	82,000	8,000
½	20	1,000	100	100,000	20,000	78,000	2,000
¼	15	800	100	80,000	12,000	74,000	(6,000)

Panel B: Naive Hedge of the Expected Quantity Required

This hedge might be put in place as part of the procurement process, since it looks only at expenses. This is at best an incomplete hedge of copper costs, since the true quantity changes with copper prices. In our example, this naive hedge actually increases risk, since increased sales mean profits are high when copper prices are high.

Probability	Unhedged ($)	"Hedge" ($)	Net ($)
¼	8,000	5,000	13,000
½	2,000	0	2,000
¼	(6,000)	(5,000)	(11,000)

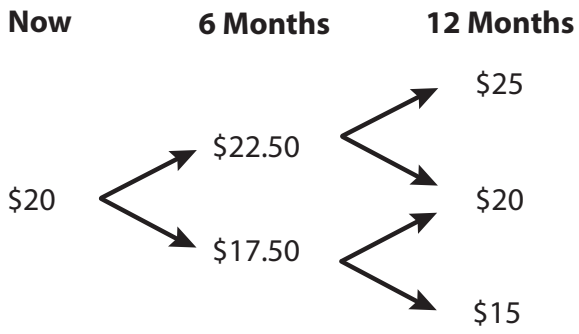
Panel C: Fully Hedged Cash Flows

A complete hedge of all the cash flows requires something more than a simple purchase of futures, since the sensitivity to copper prices of the unhedged profit or loss is higher when copper prices are low than when copper prices are high.

Probability	Unhedged ($)	Hedge ($)	Net ($)
¼	8,000	(6,500)	1,500
½	2,000	(500)	1,500
¼	(6,000)	7,500	1,500

Figure 1

Overall Price Dynamics for the Copper Hedge Example



NOTE: See text for details.

well and copper prices rise. The full hedge, the result of which is shown in Panel C, cannot be implemented by simply buying or selling copper forward one year. However, the full hedge can be implemented either by buying options or by dynamic trading in forward or futures contracts. Since this type of strategy is typical of hedging problems, it is worthwhile to derive the dynamic hedge and discuss its operation.

THE DYNAMIC HEDGE

To study the dynamic hedge, we need to understand the trading opportunities and information between now and realization of the cash flows a year from now. The sensitivity of the firm's value to copper prices varies in response to the interim information, and this changing sensitivity should be reflected in our trades.

In the current example, we assume that the firm is using copper futures contracts to hedge changes in copper prices. Futures serve the same economic purpose as forward purchases, but they are somewhat different logistically, since for futures the money changes hands immediately when the prospective value of copper rises and falls. If we buy one futures contract, then at the end of each day we are given (more literally, our margin account is credited with) the change in futures price over the day. If we sell (or *short*) one futures contract, then we must pay the change. If the futures price increases from $50 to $55, then the owner of two futures contracts will collect $10, and someone who has sold two futures contracts will have to pay $10. If the futures price instead decreases from $50 to $45, the person short two contracts collects $10, and the person long two contracts has to pay $10. In general, the futures price need not be exactly equal to the price we would pay for forward purchase, but for most purposes we can think of the two as being the same.[2]

In the actual economy, information arrives minute-by-minute and a firm can trade on copper prices almost continuously in time. For our simple example, information arrival and trading occur now, in six months, and again in a year. (This is not an essential simplification; while the

Table 2

Cash Flows from the Dynamic Hedge

Future price path ($ per unit)	Cash now	No. of contracts now	Cash in 6 months ($)	No. of contracts in 6 months	Cash in 1 year ($)	Pre-hedge cash flows ($)	Hedged cash flows ($)
20-22.50-25	0	(1,333)	(3,333)	(1,200)	(6,500)	8,000	1,500
20-22.50-20	0	(1,333)	(3,333)	(1,200)	(500)	2,000	1,500
20-17.50-20	0	(1,333)	3,333	(1,600)	(500)	2,000	1,500
20-17.50-15	0	(1,333)	3,333	(1,600)	7,500	(6,000)	1,500

analysis for a practical model requires more computations, it is conceptually no more difficult.) At the beginning of the year, the futures price of copper delivered a year from now is $20. Six months from now, the futures price will be either $22.50, with probability 1/2, or $17.50, also with probability 1/2. The overall price dynamic is shown in Figure 1. The price at a node in the tree is the price paid in a firm commitment to buy copper one year from now. From a given node, an up or down move is equally likely, with probability 1/2, so any given price path has probability $1/4 = 1/2 \times 1/2$. Consistent with Table 1, the ending node of $20 is twice as likely as the other ending nodes because it can be reached by either an up move followed by a down move (probability 1/4) or a down move followed by an up move (probability 1/4). A final price of $25 comes only from two up moves (probability 1/4), and a final price of $15 comes only from two down moves (probability 1/4).

To derive the full dynamic hedge, the firm requires one more piece of information, which is the rate at which futures gains or losses will be reinvested, which we assume to be 5 percent simple interest over six months. (Actually, the rate we choose will not affect the hedged cash flows in Panel C of Table 1, since increasing this rate will result in a completely offsetting decrease in the number of contracts we hold over the first six months.) Holding one futures contract at one node implies a gain of $2.50 (given an up move) or a loss of $2.50 (given a down move), which is reinvested until the end at the interest rate. From this we can use simple algebra to derive the solution. In the example, the full hedge is implemented by the following strategy: At the start, the firm sells $1,400/1.05 \sim 1,333$ futures at the futures price of $20. If the futures price goes down to $17.50, the firm increases the short position to 1,600 contracts, while if the futures price goes up to $22.50, the firm reduces the short position to 1,200 contracts.

The terminal cash flow generated by the hedge (including reinvestment) is analyzed in Table 2. For example, the second row shows the effects of the hedge when prices go up and then down (from $20 to $22.50 and back to $20). The hedge starts with no initial cash. It shorts 1,333 contracts; and when in six months the futures price goes up by $2.50, $2.50 × 1,333 = $3,333 is borrowed and the short futures position is reduced to 1,200 contracts. When the futures price falls by $2.50, $2.5 × 1,200 = $3,000 in profits are collected; and after payment of $3,333 × 1.05 = $3,500 on the loan, net cash from the hedge is $3,000 – $3,500, for a loss of $500. Added to the unhedged cash flow in that state of $2,000 (from Panel A of Table 1), the hedged cash flow is $1,500. The calculations in the other states work the same way.

We can see now that the dynamic hedge was chosen so that the reinvested proceeds of the hedge plus the original cash flows are made to be the same in every contingency. The necessary hedge can be computed by working backward from the end. The first two rows differ only in the price performance over the last period. Since the difference in pre-hedge cash flow for these two scenarios is \$8,000 – \$2,000 = \$6,000 and the difference in futures prices for the two scenarios is \$25 – \$20 = \$5, we require \$6,000/\$5 = 1,200 contracts to replicate the cash flows or short 1,200 contracts (the offsetting position) to hedge the cash flows. Given the calculated hedge at the last date, the calculation at the next-earlier date proceeds in the same way, and so forth back to the start. The entire strategy can be computed by looking at the linear equations implicit in Table 2 or by standard techniques described in option-pricing textbooks.

While the model underlying the hedge for the simple example probably seems too simple, it is in fact similar (except for the number of intermediate trading dates) to the binomial models used successfully in practice. Adding the additional subperiods is straightforward, given modern computing resources.

SOME FUNDAMENTAL QUESTIONS

In the example in the previous sections, we assumed that hedging is desirable. However, this assumption is far from obvious and it is useful to examine potential motives for hedging.

Why Should We Hedge?

The reason for hedging should link back to the overall objective of the firm, which is to create or enhance economic value. There is a general issue of whether the firm should maximize narrowly the value to shareholders, the total value to all financial claimants, or some more general social value to a variety of stakeholders. This distinction is not so important to us; most importantly, we assume that taxes (governments' claims) are not part of what we are optimizing, and for concreteness we speak of maximizing value to shareholders in the firm.

The first and most obvious effect of hedging is that it reduces the volatility of the value received by shareholders. Unfortunately, this does not have any value for most shareholders in a large publicly traded firm, who hold the shares in a well-diversified portfolio and for whom the additional risk is unimportant. Indeed, a conflict of interest may exist between the majority of shareholders and large shareholders (for example, members of the founding family who hold 30 percent of the shares and whose holdings are undiversified): Expending resources to reduce risk may benefit the large shareholders at the expense of the rest of the shareholders. Management may have a similar conflict, since risk threatens their jobs and they may have a significant proportion of their wealth tied up in the firm's shares. Since most shareholders in a publicly traded firm would not care about the additional risk attributable to copper price exposure, this is not a good reason for hedging. (On the other side of the equation, the cost of hedging may be very small; we consider this further in a later section on cost issues.)

A more subtle argument for managing copper price risk is that failure to do so may cause ancillary damage within the firm. As an extreme case, adverse copper price movements may

push the firm into bankruptcy, which has a number of deadweight costs to the firm, such as payments to lawyers and accountants and the loss of profitable future projects. More normally, unhedged risk exposure may tend to increase taxes, on average: While the government receives additional tax payments when the copper price move is favorable, an unfavorable move does not create a compensating tax reduction, given that tax offsets may only be deferred (and may even be lost). A related tax reason for managing copper prices is that the reduction of risk makes it possible to maintain more leverage to reduce corporate taxes and avoid "double taxation." Double taxation is the payment of both corporate and personal taxes on cash flows going to equity, compared with payment of only personal taxes on cash flows going to debt, since interest expense is an offset to income in the computation of corporate income tax. While there are no personal taxes for institutional investors—and therefore no *double* taxation—the parallel argument—single taxation versus no taxation—is valid and even more powerful for institutions. For individuals there is at least a possibility that the corporate tax on equity will be offset by lower taxes at the individual level through deferred realization of gains or by a lower capital gains rate. For tax-exempt or tax-deferred investors, the extra tax is unmitigated.

A third argument for managing copper price risk is that many firms have a policy of smoothing earnings, and hedging can reduce earnings volatility. Although this is common practice, it is hard to endorse since it seems to be an expenditure of the owners' resources to minimize the amount of information getting out to the owners. (In principle, smoothing earnings might be used to eliminate temporary variations and provide a clearer picture of long-term value, but it seems more typical that smoothing is intended to avoid quarters of poor performance without necessarily distinguishing between short- and long-term shocks.) This use of hedging may make management more comfortable and minimize criticism, but it is not in the interest of shareholders (at least in any obvious way). In some cases, hedging could be justified by the argument that it avoids restrictive debt covenants, but such covenants are far from binding in all but a small proportion of firms that smooth earnings. More common is the opposite extreme case, in which the internal objective of the firm is to ensure that earnings do not fall. Hedging for this purpose may make management comfortable—indeed too comfortable—but it discourages profitable innovation. A related strategy for limiting earnings volatility is to maintain a low level of financial leverage, which implies a large voluntary tax contribution that is not in the interest of shareholders.

A fourth argument for managing copper price risk is to make it easier to give managers incentives to produce profits: By hedging risk, we can make (for example) a division manager's compensation depend closely on value added that the manager can influence rather than what the manager cannot influence (the actual realization of copper prices). This argument for managing copper price risk implies that it may be optimal to manage copper price risk at the division level even if copper prices do not represent a significant contribution to the firm's cash flow. Of course, this strategy begs the question of why it cannot be done more cheaply (for compensation purposes only) using a paper portfolio.

Whatever the reason for hedging, we should, in principle, quantify the benefits from hedging and the trade-offs involved. This is a relatively unexplored area and more research is needed.

What Risks Should We Hedge?

The question of what risks to hedge must be subordinate to the question of why we should hedge. If there is not a compelling reason to hedge a particular source of risk, then we probably should not be hedging it. One important issue is the sense in which we would hedge a certain type of risk. For example, suppose we are hedging a bank's exposure to interest rate risk. Should we hedge the direct interest mismatch of existing assets and liabilities, or should we hedge the full economic value, which would include the value of future business? For example, a bank may find that, as interest rates rise, core deposits tend to be lost. Current accounting methods make it hard to hedge this sort of risk without penalty (and the risk-based capital requirements from the Basel I and II Accords penalize almost all hedging because capital must be increased once for the underlying cash flow and once again for the hedge[3]). There is also a related question of whether to hedge cash flows or value. In practice, hedging cash flows out a year is much different from hedging the firm's entire value. If the purpose of hedging is to eliminate sources of noise beyond the manager's control, it may even be appropriate to hedge particular accounting numbers used in computing compensation rather than hedging cash flow or economic value.

With What Instruments Should We Hedge?

Firms have many instruments they can use to hedge common risks, such as exposure to interest rates, foreign exchange rates, and commodity prices. For example, to hedge U.S. interest rates we can use bonds, repurchase agreements, Treasury bond futures, swaps, caps, or collars. The choice among this set would be determined by pricing and transaction costs, adequately matching the tool to hedging needs, and accounting implications.

Support Your Investment Banker

A common approach of managers planning to hedge is to turn the whole problem over to an investment banker who, after all, has the expertise and the traders who can put the hedge in place and is happy to provide "free" advice on what to do. As in all markets, the "free" advice is priced out in what you pay for the hedge, and then some. To avoid paying too much, it is best to understand how the hedge works and how much it should cost. Ideally, such expertise should be available in-shop; otherwise, it is worth the expense of hiring an expert to monitor the prices being paid to the investment banker. In general, competition among investment bankers may be useful in reducing the cost, but competition will not necessarily produce any incentive to tell the client when hedging is unnecessary.

ACCOUNTANT: FRIEND OR FOE?

Suppose we put in place the optimal hedge computed above, using the model for demand and option-pricing theory to determine the correct holding in futures to offset the risk in the cash flows. What will this do to our accounting statements? In general, accounting looks at the present and the past: Accountants favor methods with easily replicable results, especially since standard mechanical rules, even if inaccurate, are easy to defend if the firm has followed certain

accounting standards such as the Generally Accepted Accounting Principles or GAAP in the United States. Hedge accounting is a relatively new and technical area, and the accounting profession is only starting to address the important issues involved.

First, what the well-thought-out (dynamic) hedge transactions do to our accounting statements depends critically on whether these transactions qualify for the so-called hedge accounting treatment according to official pronouncement by the Financial Accounting Standards Board (FASB) or International Financial Reporting Standards (IFRS). The current guide for hedge accounting standards in the United States is the FASB's Statement of Financial Accounting Standard (SFAS) No. 133: *Accounting for Derivative Instruments and Hedging Activities*.[4] Even if the transactions qualify for "hedge accounting," effects on financial statements differ depending on which one of the two allowable hedge accounting methods applies: fair value hedge or cash flow hedge.[5]

Briefly, a fair value hedge refers to companies entering into derivative contracts to hedge the fluctuation of the fair value of existing assets, liabilities, or commitments otherwise not recognized on the company's balance sheet. A cash flow hedge refers to companies entering into derivative contracts to hedge the fluctuation of their future cash inflow or outflow. For transactions in our example, the economic risk to be hedged is the fluctuation of a company's future sales (cash inflow) and production costs (cash outflow). In hedge accounting terms, these are "forecasted future transactions," so the proper hedge accounting, if applicable, would be so-called cash flow hedge accounting. (See the glossary of hedge accounting terms in the appendix.)

Several conditions must be met to qualify for cash flow hedge accounting treatment[6]: (i) The nature of the risk to be hedged must be interest rate risk, price risk, foreign currency exchange rate risk, or credit risk. (ii) The hedge must be deemed effective. (iii) Formal and complete documentation of the hedging activities must be prepared. (iv) The forecasted future transactions must be either a single transaction or a group of similar transactions with an unrelated counterparty (i.e., not a subsidiary or affiliated company). And (v) the likelihood of the forecasted future transactions occurring must be "probable."

The hedge in our example presents an interesting accounting challenge. The source risk in our example can be thought of as a price risk (i.e., the price of copper fluctuates between $15 and $25), which is an allowable risk for hedging under SFAS 133. Economically equivalent, the source risk in our example can be thought of as quantity risk (i.e., the quantity of the total sale fluctuates between 800 and 1,200 contracts), which is not an allowable risk for hedging treatment. Supposing these transactions survive all hedge accounting tests, the transactions' effects on the company's financial statements are generally twofold. First, on the quarter-end or year-end balance sheets, the fair value of the future contracts must correspond to their market value. Second, prior to the settlement of the futures, unrealized gains or losses on the contracts will be reflected directly in shareholders' equity without going through the company's income statement—that is, the unrealized gains or losses do not affect, or "bypass," net income. At maturity, all unrealized gains and losses are reclassified as an earnings component into the net income of the maturity fiscal year. In other words, prior to maturity, the accumulated unrealized gains and losses resulting from cash flow hedges are temporarily stored in a special place (a shareholder equity account called "accumulated other comprehensive income" [AOCI]) before being

"recycled" into earnings at termination. The presumed reason is that since the forecasted future transactions to be hedged are not reflected on the balance sheet prior to maturity (thus their gains and losses are necessarily not reflected in the income statements prior to maturity), flowing the gains and losses from the futures into the income statement alone does not makes sense.

However, there is a catch. SFAS 133 rules allow only the "effective" portion of the cash flow hedge to be stored in AOCI prior to maturity. The effective portion of the hedge refers to the changes in the expected cash flows of the forecasted future transactions to be hedged. Any "ineffective" portion of the cash flow hedge must be reported as part of net income as incurred. Our example poses a subtle challenge here as well. Suppose the hedge risk is deemed as the price risk of copper and the forecasted future transactions are the future purchases of copper; then what is deemed effective for the hedged item (the total cost of a raw material) may not be the real economically effective hedged total risk to the company's net cash flow.

In fact, if the ineffective portion becomes too big such that the so-called delta ratio falls outside the 0.80 to 1.25 interval (commonly referred to as the "80/125" standard in practice), the hedge may be disqualified from receiving the cash flow hedge accounting treatment. If that is the case, all gains and losses (past, current, and future) will be reflected in earnings. The same accounting treatment is applied to any derivative holdings that fail to qualify as SFAS 133 hedge accounting.

Failure to qualify as a hedge often penalizes true economic hedging. An unqualified hedge typically reduces volatility of future cash flows but increases volatility of reported earnings. This volatility is especially damaging when it causes a violation of debt covenants or capital requirements imposed by regulators. Earnings volatility may also subject management to criticism; given the current hysteria over derivatives, we may want to pardon a manager who forgoes an economically useful hedge to avoid the appearance of "risky exposure to derivatives." Part of the problem is that there seems to be no simple test, given the current state of hedge accounting, that the lay public can apply to distinguish risky speculation from good hedging.

One interesting feature of the accounting rules is that hedges that are economically equivalent may have very different accounting treatments. Suppose in the example above that demand does not depend on copper prices (putting the same number in all of the "Units sold" column in Panel A of Table 1) and we are simply interested in hedging the input cost at expected demand. Then it might seem equivalent to hedge by (i) having a long-term contract with a supplier, (ii) buying copper futures, or (iii) buying shares in a company whose share price tracks copper closely. However, (i) the contract with a supplier has no impact on earnings before the actual sale unless the company decides to hedge the fair value risk for noncancelable, fixed-price, long-term supply contracts; (ii) buying copper futures is covered by SFAS 133 as discussed above; and (iii) shares in the copper company are accounted at fair value, but unrealized gains and losses may or may not appear in earnings (SFAS 115). This approach is based on a general recognition that current reporting practice is often misleading and a paucity of good ideas on how to patch things up. It seems that hedging tends to magnify the problems inherent in the accounting profession's tension between historical cost and mark-to-market cost. Further, the current hedging SFAS 133, while superseding prior vague standards (SFAS 80, 105, and 119, among others), was itself so vague initially that its effective date was delayed and has been amended quite a few times;

it has become incredibly detailed since then, given the voluminous implementation guides issued by the FASB staff. It is arguable that the cost of compliance with SFAS 133 can become a nontrivial issue for many companies.

It should be mentioned here that some people have proposed universal adoption of mark-to-market (or intrinsic value) accounting, which is "obviously" the correct thing to do because that is a good estimate of what the firm is actually worth, and any hedge would be seen for what it is. Unfortunately, it is not at all clear what this means. For example, do we include future sales in our valuation? If so, how far in the future do we go and how do we forecast and value the future flows? Anyone who has been involved in capital budgeting knows that estimates of future cash flows are often inaccurate and may reflect the forecaster's optimism more than the prospects for the firm. Even without these conceptual problems, introducing a whole new system of accounting is not a trivial matter. While we note that current accounting standards are deficient for measuring risk, we do not claim that it is easy to do better. The differences in accounting treatments of economically equivalent hedges may allow firms to hedge despite the deficiencies in the accounting standards. Regardless of whether a firm that is hedging properly can avoid looking bad, it is clear that a firm that is not hedging at all—or even increasing risk—can look fine.

COST ISSUES

What is the cost of hedging? It is tempting to think that the cost of the hedge is the cost of any securities purchased in the hedge program. In fact, the hedge is often bundled with an investment. It is a fair investment to buy a call option for its intrinsic value, and absent market imperfections there is no cost in doing so. In practice, the cost includes transaction costs such as commissions, bid-ask spread, and any internal costs of trading (e.g., hiring a trader and setting up accounting oversight). For publicly traded contracts in liquid markets, the costs are probably small and easy to measure. When hedging uses custom contracts provided by investment bankers, the costs are hard to assess (because they are built into pricing) and may be much larger.[7] On a more esoteric point, we may also want to include in the cost of hedging the alternative use of any capital tied up in the investment or in margin or variation accounts. On another subtle point, a hedge may be more costly than it appears if its pricing and tax treatment make it inappropriate for the firm.

What is the marginal cost that should be used as an input for decisions about pricing the output? It is probably common to use the hedged price, but in fact the marginal cost of the commodity at the time of use is the spot market price (assuming an active market that was probably necessary to implement the hedge in the first place). It is irrelevant that the price has been locked in for a fixed quantity, since that is sunk, and the profit will be collected (or the loss borne) on the hedged quantity regardless of how much or little is actually used. If more is needed, the shortfall will be purchased at the spot price. If less is needed, the excess will be sold at the spot price. In either case, the marginal cost is the spot price. If the marginal cost is taken to be the hedge price (or some average price), value may be discarded. For example, suppose the spot price is higher than the hedge price. Then a computation assuming that marginal cost equals the average price or the hedged price would understate the true cost of buying more of the input, and addi-

tional units could be sold when it is more profitable to sell what can be produced from the hedged quantity of inputs.

What is the transfer price that should be used when the commodity is procured by one unit in the firm and used by another? For accounting purposes, the organization should decide upfront how profits and losses in the hedging program will be shared. It is probably best to plan to do so in a way that hedges cash flows in each unit, since that approach ties compensation in each unit more directly to performance within the manager's influence. An inherent unfairness may result if the sharing of the hedge's profit and loss is not decided in advance. For example, suppose the transfer price is ambiguous or renegotiable. If the transfer price is the market price when the market price is low but a hedged price when the market is high, the purchasing unit gets a "free option" and the procuring unit loses—regardless of whether it is hedged. The free option allows the unit to buy at the hedged price or the market price, whichever is less. The selling unit always loses money.

RISK MANAGEMENT POLICY

Given that standard accounting procedures do not provide a particularly useful picture of the quality of a firm's hedging program, it is especially important for management to adopt and implement an understandable and effective risk management policy. Such a policy should specify the goal and scope of any hedging activity, and it should dictate the degree of centralization and the control systems. Furthermore, the policy should provide for oversight and evaluation of the effectiveness of hedging.

A common feature of the large publicized trading losses is a failure of control systems. Financial firms face a particular temptation to have inadequate controls. Because firms want to retain successful traders, they may tend to be sympathetic to traders' insistence that the bureaucracy should not interfere with their work. A failure to separate the operations and accounting functions from trading was an essential common thread in the losses of over a billion dollars each at Barings, Daiwa Bank, and Sumitomo. In each case, the loss was attributable to a single trader. It is important to devote serious talent to the job of monitoring traders, even though the monitoring job is less glamorous, somewhat unpleasant, and, when things are going well, seemingly unproductive.

Besides the scenario of speculation under the guise of risk management, risk management can be counterproductive if it is too localized. To illustrate, the example we discussed earlier showed how a procurement department that is hedging material costs may actually make overall cash flows more variable if input prices tend to be high when the industry does well. Less damaging, but probably still wasteful, is the practice in which companies use different parts of the firm to offset hedging or they hedge economically irrelevant risks (such as risks that represent an insignificant part of a firm's cash flow volatility). For most firms, the benefits of centralization (better control, economies of scale, and cost saving resulting from internal netting) will outweigh the costs (mostly the difficulty of communicating and aggregating needs). Of course, it is a good idea to have a formal policy in either case, whether risk management is centralized or dispersed. Unfortunately, there is a tension between the economic goal of having a centralized

overall hedge and the accounting requirement in which a qualified hedge can apply only to a specific transaction.

A good risk management policy should state the goals of the hedging program. Is it the firm's policy to hedge the value of the firm or, alternatively, earnings or dividends paid to shareholders—and if so, what risks should be hedged and what risks should be borne by the shareholders? Should hedging be implemented on a divisional or departmental level (to improve planning and incentive compensation) when that hedging does not reduce the overall variability of the firm's value? Should the hedging program focus on cash flows, earnings, tax avoidance, or something else? We do not yet have definitive answers to these questions, but at least a consistent policy will minimize offsetting efforts.

One important (but probably often neglected) aspect of a risk management program is the need for ex post evaluation. Especially because these programs are relatively new, it is entirely possible to design a program that is ineffective or that even increases risk (such as the naive hedging strategy in our copper price hedging example). Only retrospective analysis of the results can verify that the program is actually reducing risk. The retrospective analysis should also look at any side effects of the hedging—for example, variation or margin account payments required to maintain the hedge.

Finally, even in the absence of hedging, the risk management policy should monitor and limit the amount of risk taken, especially through the impact of decisions made unilaterally and/or at a low level in the firm. In capital budgeting, small projects within a department's budget can be undertaken without outside evaluation, but very big projects have to be "kicked upstairs" for approval. For many types of activities, the possible loss is normally bounded by the capital outlay so that this process tends to limit the amount of damage done by unilateral decisions at a low level. However, for some activities (such as trading futures), only a small capital outlay is required to create a large liability. Practitioners have developed the notion of a "risk budget" that says how much risk a trader or unit of the firm is allowed to take. Just like capital budgets, risk budgets do not have to be absolute constraints, and in practice they specify how big a decision the individual or unit can make without bringing other people into the decision-making process. This is a developing area: The measures of risk used (e.g., value at risk) are ad hoc and are not motivated by economic fundamentals, but they are still better than nothing.

CONCLUSION

Risk management is an important and difficult area of corporate policy. We have seen news accounts of disastrous failures in risk management. Less spectacular, but perhaps more important, is the widespread use of futures contracts and swaps to hedge foreign exchange, interest rate, and commodity risks, since without this ability to hedge, many profitable businesses would be too risky.

The next few years should be especially interesting[8] as companies work on implementing vague new accounting standards that require them to describe their risk exposure. Now is also an exciting time for the development of internal controls and policies as companies work on developing effective hedges while avoiding catastrophic losses.

NOTES

1 Cash flow is the accounting notion of actual cash coming into or going out from operations. Unlike profits, cash flow does not include depreciation or amortization, but it does include (as an offset) investment in capital. In our examples later, we treat the two the same, although this is not appropriate except in the case of very simple businesses that rent any required capital.

2 In fact, if interest rates are nonrandom (so reinvestment rates are known in advance), the absence of arbitrage implies that the forward price must equal the futures price, although a futures contract has more impact since the change in value in a futures contract is received up-front, while the change in value in a forward contract occurs at maturity.

3 It is too early to tell what effect, if any, the Basel III Accord, passed in late 2010, will have on the status quo because each individual country is in the process of specifying the local implementation of the accord.

4 This statement was issued in 1998 but has since been heavily amended by subsequent follow-up standards, such as SFAS 138 in 2000, SFAS 149 in 2003, SFAS 155 in 2006, and SFAS 161 in 2008.

5 There is a third method called a net investment hedge, which does not apply here.

6 To qualify for any hedge accounting variety, the hedge instrument must be a derivative instrument in the eyes of SFAS 133, which defines derivatives as instruments with the following three characteristics: (i) having one or more underlying prices or values and one or more notional amounts or payment provisions; (ii) requiring no or a relatively small initial net investment; and (iii) requiring net settlements by a cash payment between parties or settlements by delivery of an asset that can easily be converted to cash or is another derivative.

7 In fact, there are recent high-profile cases where major private colleges such as Harvard incurred millions of dollars in costs to terminate interest rate swap agreements; see Lauerman and McDonald (2009).

8 This assertion in the original article was definitely an understatement, given the role of risk management failures in the financial crisis of 2007-09. And the current debates on hedge accounting imply that the same forecast is likely to be true today.

REFERENCES

Black, Fischer and Scholes, Myron. "The Pricing of Options and Corporate Liabilities" *Journal of Political Economy*, May-June 1973, *81*(3), pp. 637-54.

Cox, John A.; Ross, Stephen A. and Rubinstein, Mark. "Option Pricing: A Simplified Approach." *Journal of Financial Economics*, September 1979, *7*(3), pp. 229-63.

Lauerman, John and McDonald, Michael. "Harvard's Bet on Interest Rate Rise Cost $500 Million to Exit." Bloomberg.com, October 17, 2009; http://www.bloomberg.com/apps/news?pid=newsarchive&sid=aHou7iMlBMN8.

Rubinstein, Mark and Leland, Hayne E. "Replicating Options with Positions in Stock and Cash." *Financial Analysts Journal*, July/August 1981, *37*(4), pp. 63-72.

FURTHER READING

Risk Management Information

For general information on risk management, some banks have issued guides that may be useful. For example, the *J. P. Morgan/Arthur Anderson Guide to Corporate Risk Management* (London: Risk Publications, 1997) is a primer on risk management, while *The Chase Guide to Risk Management* is an extended glossary published by Chase Manhattan Bank (1993) in association with *Risk* magazine. Another good general resource is "A Survey of Corporate Risk Management" by Matthew Bishop (*Economist*, Special Edition, February 10, 1996, pp. 1-22). A more recent overview is *Risk Management: A Review*, by Sébestien Lleo (London: CFA Institute, 2009).

Hedging with Options

Some observers have debated the advisability of hedging with options. One criticism is that hedging with derivatives often amounts to gambling with firm money at the encouragement of banks (see Ralfe, John. "Betting Your Hedges." *Risk*, July 1994, *7*(7) pp. 22-23). One good point is that customized option positions are unlikely to be a good value, especially if it might be necessary to unwind the position before maturity (see Westby, D. "Caveat Emptor." *Risk*, June 1995, pp. 24-25). A different argument is that poorly constructed strategies with poor disclosure can lead to legal troubles for managers and directors (see Falloon, W. "Courting Trouble." Falloon, *Risk*, August 1994, *7*(8), pp. 32-33). On the other side of the debate is the suggestion that hedging with options is a good idea and that most criticisms are unjustified (see Schewitz, M. "Keep Those Options Open." *Risk*, October 1995, *8*(10), pp. 35-36).

Hedging Programs That Failed

There are, of course, many articles in the popular press citing specific hedging programs gone badly. The case of Metallgesellschaft, for example, sparked a debate among academics because it may be a case of miscalculation of the hedge. In particular, for commodities subject to temporary shortages, there is no reason to believe that price uncertainty over longer horizons can be hedged effectively by available traded options at short horizons, although this is often assumed to be the case. Several scholars have studied this question using theoretical tools (see Ross, Stephen A. "Hedging Long Run Commitments: Exercises in Incomplete Market Pricing." *Economic Notes by Banca Monte*, 1997, *26*(2), pp. 99-132); empirical tools (see Edwards, Franklin R. and Canter, Michael S. "The Collapse of Metallgesellschaft: Unhedgeable Risks, Poor Hedging Strategy, or Just Bad Luck?" *Journal of Applied Corporate Finance*, Spring 1995, *8*(1) pp. 86-105, or Bakshi, Gurdip; Cao, Charles and Chen, Zhiwu. "Pricing and Hedging Long-Term Options." *Journal of Econometrics*, January 2000, *94*(1-2), pp. 277-318); or both theoretical and empirical tools (see Pirrong, Stephen C. "Metallgesellschaft: A Prudent Hedger Ruined, or a Wildcatter on NYMEX?" *Journal of Futures Markets*, August 1997, *17*(5), pp. 543-78,). The defense of Metallgesellschaft's hedging program is that it was a textbook hedge that would have done fine if not interrupted (see Culp, Christopher L. and Miller, Merton H. "Hedging a Flow of Commodity Deliveries with Futures: Lesson from Metallgesellschaft." *Derivatives Quarterly*, Fall 1994, *1*(1), pp. 7-15). Other high-profile cases include Orange County (in 1994), Long-Term Capital Management (in 1998), Amaranth (in 2006), and quite a few cases of "rogue traders" within large financial institutions.

Value at Risk

Value at risk is one methodology widely used to quantify various common sources of financial risk (see Duffie, Darrell and Pan, Jun. "An Overview of Value at Risk." *Journal of Derivatives*, Spring 1997, *4*(3), pp. 7-49). This methodology has its critics, both because an objective measure is difficult to agree on (see Beder, Tanya S. "VAR: Seductive but Dangerous." *Financial Analysts Journal*, September/October 1995, *51*(5), pp. 12-24) and because value at risk neglects idiosyncratic risk and some market sources of risk. An updated and readable reference is Phillippe Jorion's *Value-at-Risk: The New Benchmark for Managing Financial Risk* (Third Edition. New York: McGraw-Hill, 2007).

Hedge Accounting Sources

Mark Trombley's *Accounting for Derivatives and Hedging* (New York: McGraw-Hill, 2003) is a good introduction to the accounting mechanics on hedge accounting as specified by SFAS 133. Stephen Ryan's *Financial Instruments and Institutions: Accounting and Disclosure Rules* (Second Edition. Hoboken, NJ: John Wiley & Sons, 2007) is a good source of accounting and disclosure issues related to banks and other financial service entities.

APPENDIX

A Short Glossary of Risk Management Terms

Binomial model. The binomial option-pricing model, developed by Cox, Ross, and Rubinstein in 1979, assumes that the stock return over a short time interval has one of two values. The binomial model is a popular alternative to the Black-Scholes model because it is flexible and easy to implement on a computer.

Black-Scholes model. This is the original modern option-pricing model (see the boxed insert on p. 275).

Call option. A call option is a contract that gives the owner the right to purchase a share of the underlying asset in exchange for the contractually specified strike price (or exercise price). An *American* call option can be exercised at any time before maturity, while a *European* call option can be exercised only on the maturity date.

Cap. An interest rate cap is a promise to pay the excess of an interest rate above some level in each of a number of periods. Caps are useful for containing the risk of rising borrowing costs.

Collar. A collar combines the cash flows of buying a cap and selling a floor. It is useful for containing the risk of rising interest rates (like a cap); including the floor gives up some profit potential when rates fall to help to pay for the cap.

Floor. An interest rate floor is a promise to pay the shortfall of an interest rate below some level in each of a number of periods. Floors are useful for locking in a minimum return.

Forward contract. A forward contract gives the owner the right and the obligation to buy a specified amount of a commodity at a specified price at some specified date in the future.

Futures contract. A futures contract is similar to a forward contract except that there is daily settlement (i.e., each day the parties to the contract exchange money representing the market-determined change in value of the contract). Daily settlement minimizes the need for credit checks and large margin accounts (which are held as collateral), since only one day's price variation is at risk.

Hedge. Hedging a position (or entering a hedge) is undertaking another activity with offsetting risk. Some common hedging instruments include insurance, futures contracts, and options.

Long position. To take a long position (or to "be long") is to purchase an asset or futures.

Put option. A put option is a contract that gives the owner the right to sell a share of the underlying asset in exchange for the contractually specified strike price (or exercise price). An *American* put option can be exercised at any time before maturity, while a *European* put option can be exercised only on the maturity date.

Short position. To take a short position (or to "sell short") is to assume the opposite of a long position. In the case of futures, they are simply sold in the market. Shares and other securities are borrowed (for a nominal fee) and then sold in the market, with the promise of buying some shares later to return the borrowed shares. In the meantime, the short must pay any dividends or coupons that are due the person from whom the shares were borrowed. The cash flows for a short position are the negative of the cash flows for a long position.

Value at risk. Value at risk (VAR) of a portfolio is typically defined as the maximum loss within a confidence level of $1-\alpha$ (e.g., if $\alpha = 5$ percent, then a 95 percent confidence level) that the portfolio could incur over a specified time period (e.g., three days). Formally, it is defined as $VAR(X; \alpha) = -\{X|F(X) < \alpha\}$, where $F(.)$ is the cumulative density function of portfolio value over the specified period.

A Short Glossary of Hedge Accounting Terms

Accumulated Other Comprehensive Income (AOCI). In hedge accounting, having an entry in AOCI is used as a technical device for breaking the typical accounting articulation between the balance sheet and income statement and, as a result, having changes in the asset or liability on the balance sheet that do not flow through from the income statement. For a given year, these changes are called other comprehensive income. AOCI is the cumulative amount of other comprehensive income shown as a component of stockholders' equity.

Basel Accords. The Basel Accords are a set of agreements set by the Basel Committee on Banking Supervision, which provides recommendations on banking regulations in regard to capital risk, market risk, and operational risk. The first accord was the Basel I issued in 1988. Basel II was published in June 2004 and its implementation was interrupted by the 2007 financial turmoil. A post-crisis agreement, Basel III, was reached in 2009. The Basel Accords were intended to keep banks safe and create a level playing field, but they have failed on both counts. As learned from the financial crisis, they have not kept the banks safe and national differences in implementation have prevented the accords from creating a level playing field.

Cash flow hedge. A cash flow hedge is a hedge of the variability of cash flows related to a future transaction. The hedged item might be (i) a contractual cash flow related to an existing asset or liability or (ii) a forecasted transaction that is expected to occur but is not yet subject to a contract.

Fair value. Fair value is the amount at which an asset or liability could be bought, sold, or settled in a voluntary transaction between unrelated parties, the best evidence of which comes first from quoted market prices from active exchanges and then dealer quotes for items not actively traded or traded between principals rather than on an exchange. In the absence of price quotes, fair value must be estimated using pricing models, discounted cash flows, or other techniques.

Fair value hedge. A fair value hedge is a hedge of a risk exposure pertaining to either (i) a recognized asset or liability or (ii) an unrecognized firm commitment. A recognized asset or liability is one recorded on the balance sheet—for example, inventory, accounts payable, or debt obligations. An unrecognized firm commitment is a contractual obligation that is not yet reflected on the balance sheet.

Financial Accounting Standards Board (FASB). The FASB is a private, not-for-profit organization whose primary purpose is to develop Generally Accepted Accounting Principles within the United States in the public's interest.

Forecasted transactions. Forecasted transactions are the transactions expected to occur in the future but not yet subject to a legally enforceable contract.

Generally Accepted Accounting Principles (GAAP). The set of GAAP is a codification of how certified public accountant firms and corporations prepare and present their business income and expense, assets, and liabilities on their financial statements.

Hedge accounting. Hedge accounting is an accounting approach that allows a gain or loss on a hedging instrument to be recognized in the same period that the offsetting loss or gain on the hedged item is recognized.

Hedge effectiveness. Hedge effectiveness is the extent to which changes in a hedged item's value are offset by changes in the value of the derivatives used for hedging. If the changes offset exactly, the hedge is said to be perfectly effective, and if they substantially offset, the hedge is said to be highly effective.

International Accounting Standards Board (IASB). A private accounting standard-setting body based in London, England. The board promulgates International Financial Reporting Standards (IFRS), which have been adopted by many countries around the world, including all countries in Europe. Its mission has been to offer an internationally consistent set of standards to replace possibly different domestic standards in various countries; however, it still faces many challenges ranging from different implementation choices by adopting countries to winning adoptions by large some economies, notably the United States.

International Financial Reporting Standards (IFRS). The IFRS state how particular types of transactions and other events should be reported in financial statements. The IFRS are issued by the International Accounting Standards Board.

SFAS 133. SFAS 133 is the FASB's Accounting Standard for Derivative Instruments and Hedging Activities. It requires all derivative instruments to be shown on the balance sheet at fair market value with the accounting for changes in fair value depending on the purpose of the derivative.

Why Are U.S. Firms Holding So Much Cash? An Exploration of Cross-Sectional Variation

Juan M. Sánchez and Emircan Yurdagül

Currently U.S. firms hold record amounts of cash. The authors explore cross-sectional variation in cash holdings of U.S. publicly traded firms to shed light on the reasons for this recent trend. First, they identify factors that correlate with cash holdings and then examine the evolution of these factors over the past decade. Several factors, including research and development expenditures and idiosyncratic uncertainty, are important in accounting for cross-sectional differences in cash holdings. However, these factors do not increase over time as cash holdings do; thus, it seems unlikely that they underlie the increase in cash holdings. Aggregate uncertainty, however, has recently reached record levels. This uncertainty, combined with the fact that (idiosyncratic) uncertainty correlates well with cash holdings in the cross section of firms, suggests aggregate uncertainty may be an important factor accounting for the recent trend in increased cash holdings. (JEL E43, E44, E62, G33)

Federal Reserve Bank of St. Louis *Review*, July/August 2013, *95*(4), pp. 293-325.

Currently U.S. corporations have record-high cash holdings. Many argue that this phenomenon is related to the sluggish recovery of the economy: Firms holding more cash are investing less, and this prevents the economy from taking off. While referring to the cash holdings of Apple, the president of a business association stated "Why wasn't Apple spending that money on expansion, new products and jobs? The answer is uncertainty—uncertainty over new taxes" (Brunell, 2011). The concern about cash holdings is also present in the academic literature, where the rise in cash holdings of firms has been associated with several factors, including many that started more than a decade ago. This article evaluates the role played by potential firm-level determinants cited in the academic literature by examining the cross section of publicly traded firms. In addition, it provides some evidence suggesting cash holdings may relate to aggregate uncertainty.

In theory, transaction costs and precautionary motives are the main reasons firms hold cash. On the one hand, transaction costs associated with liquidating certain assets make cash preferable in response to movements in the market. Thus, holding cash can be important because

Juan M. Sánchez is an economist at the Federal Reserve Bank of St. Louis. Emircan Yurdagül is a technical research associate at the Federal Reserve Bank of St. Louis and a PhD student at Washington University in St. Louis.

it facilitates firms' responses to profitable opportunities. On the other hand, cash holdings are desirable as a cushion to soften adverse movements in the market, especially for firms with borrowing constraints.

If a factor is to be considered important in explaining the rise in cash holdings, it must satisfy the following two characteristics: (i) It must correlate with cash holdings in the cross section of firms and (ii) it must have increased over the period during which cash holdings increased. Thus, to evaluate each factor, we construct one variable that captures differences in that factor across firms and study the cross-sectional correlation of that variable with cash holdings and its evolution over the past several years. The goal of this analysis is to identify factors that can potentially do well in accounting for the recent increase in cash holdings. Factors that correlate with cash holdings in the cross section of firms and that increased during the past decade are actually good candidates. Of course, correlation does not imply causation, and the analysis here is just a first step in trying to understand the recent increase in cash holdings.

We use Compustat data that cover publicly traded companies. This dataset includes detailed annual balance sheet information, enabling the tracking of various activities and conditions for each firm. However, since the dataset does not cover privately held companies, the patterns identified in this article do not necessarily reflect the behavior of all U.S. firms.

The article is organized as follows. The next section shows the pattern of cash holdings through time in the U.S. economy and the cash-holding behavior in the cross sections of U.S. firms. In the following section, we examine various other indicators of firms' actions and conditions and study how these variables, defined at the individual firm level, relate to the cash holding of corporations in our sample. We then look for explanations in the aggregate trends of productivity volatility and policy uncertainty and follow with our conclusion.[1]

CASH HOLDINGS OF U.S. FIRMS

In this section, we describe various facts regarding the cash-holding behavior of U.S. firms. Our main focus is on the increasing weight of cash and equivalently liquid assets on firms' balance sheets. We first describe the time pattern of cash holdings. We then study the cross section of firms and describe the distribution in the overall sample and within various subgroups.

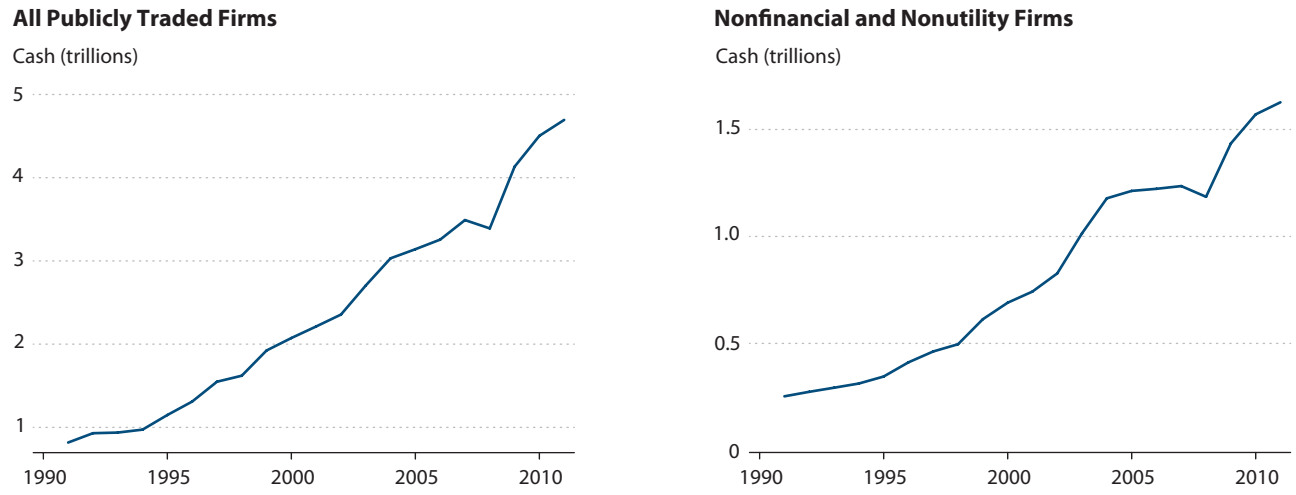
Recent Trends in Cash Holdings

A persistent and increasing pattern in cash holdings since the 1980s is notable in the aggregate behavior of U.S. firms (Figure 1, left panel). However, the pattern is much steeper after the mid-1990s. Specifically, the annual growth rate from 1979 to 1995 was around 6.5 percent, while after that point until 2011 the growth rate was above 9 percent. Stated differently, by 2011 U.S. firms were holding 4 times as much cash as they were holding in 1995 and 11 times as much as they were holding in 1979.

In the rest of the paper, we exclude financial companies and utility companies, as is standard in the literature. We do so because these groups of firms hold cash for different reasons, including regulations. After this exclusion, the scale of the aggregate cash holdings changes significantly,

Figure 1

Aggregate Cash by U.S. Firms



SOURCE: Compustat.

but the trend mentioned earlier changes only slightly (see Figure 1, right panel). In particular, from 1979 to 1995 the annual growth rate of cash holdings was close to 7 percent, while from 1995 to 2011 it was as high as 10 percent.

Although the trend in the aggregate levels of cash holdings is striking, in part this may simply be due to economic growth or inflation. The fraction of assets in the form of cash is a better measure that can be used to make healthier conclusions. Henceforth, we focus on the cash ratio, which is defined as the ratio of cash and equivalents to total assets:

$$CashRatio_{it} = \frac{Cash_{it}}{Assets_{it}},$$

where subindex i denotes the individual firm and t the period.

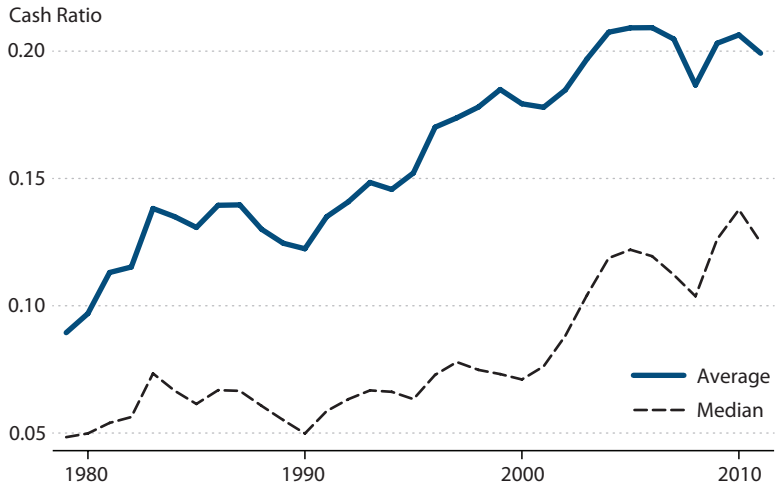
Figure 2 shows the trend of this normalized variable of cash holdings. Even though it is not as monotone as for the aggregates, the cash ratio increased substantially over the past three decades. The median of the cash ratio across firms increased at an annual rate of 4 percent from 1990 to 2000. After 2000, the trend is even steeper, until 2005. In particular, the annual growth rate from 2000 to 2005 exceeded 11 percent. This high point was followed by a drop until 2008 in an interval with the crisis, after which the cash ratio recovered very quickly. Overall, the cash ratio increased significantly in the 2000s: The 2010 ratio was almost twice as large as in 2000.

Cross-Sectional Cash Holdings

Previously we analyzed the evolution of cash holdings over time. In this section, we study how the distribution of the cash ratio is formed in the sample and how it varies between different

Figure 2

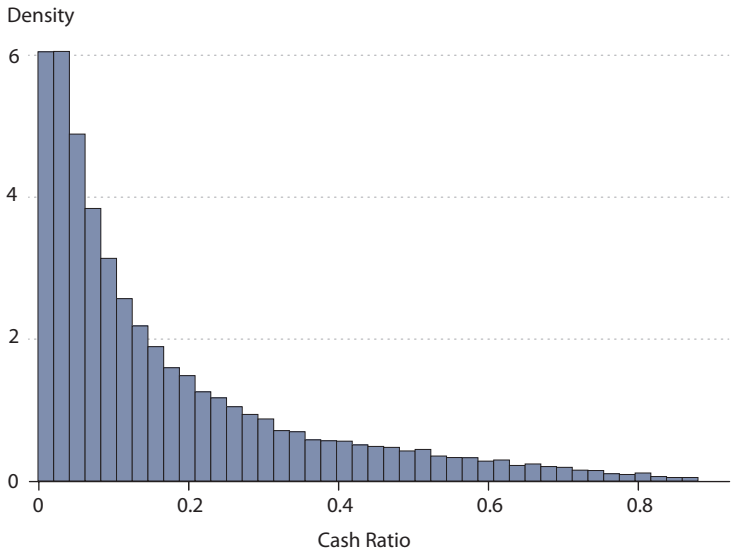
Cash Ratio (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

Figure 3

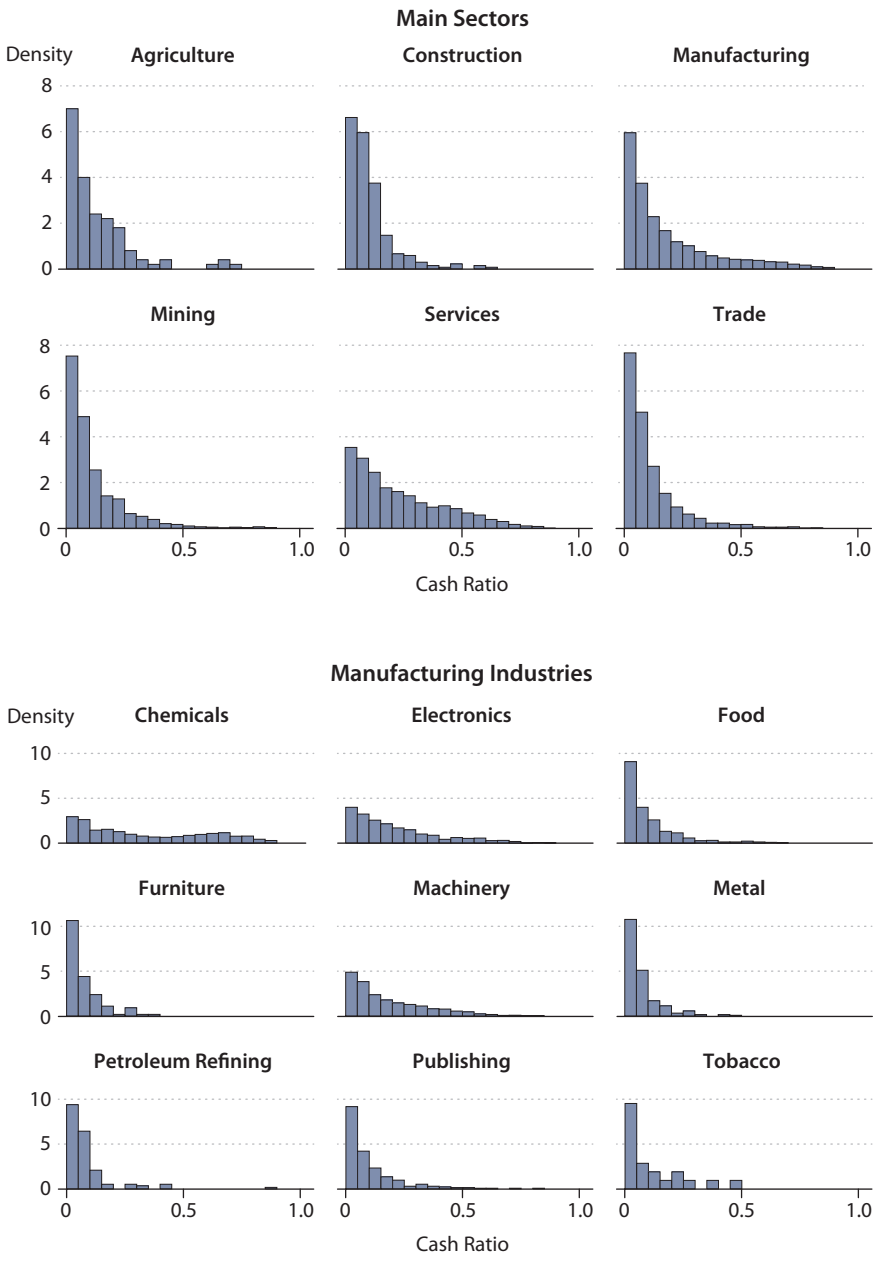
Distribution of Cash Ratio (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

Figure 4

Distribution of Cash Ratio by Sectors



SOURCE: Compustat.

groups of firms. A good approach to accomplish this and to abstract from the time dimension is to determine the mean cash ratio for each firm across time. Later we follow this approach to study other variables in cross sections of firms.

Figure 3 shows the cash ratio distribution in our sample. One observation is the fat right tail of the distribution, as there are many firms with cash ratios very close to zero and few firms with large cash holdings. In particular, 35 percent of the firms have an average cash ratio larger than 0.17, which is the mean ratio in the sample. Figure 4 shows the cash ratio distribution for various broad sectors and industries in the economy. The sector with the most disperse distribution appears to be services. As a result, services has the highest mean cash ratio among these sectors: 0.23. The sector with the next-highest ratio is manufacturing, with a mean cash ratio of 0.18. In line with these high cash ratios, these two sectors also have the highest standard deviation (SD) within the sectors in the top panel—both with 0.19—while the overall sample has an SD of 0.18.

There is also considerable heterogeneity within manufacturing; chemicals, electronics, and machinery have the flattest distributions and the highest means. Respectively, their average cash ratios are 0.33, 0.21, and 0.18 and SDs are 0.26, 0.18, and 0.16.

THE ROLE OF IDIOSYNCRATIC FACTORS

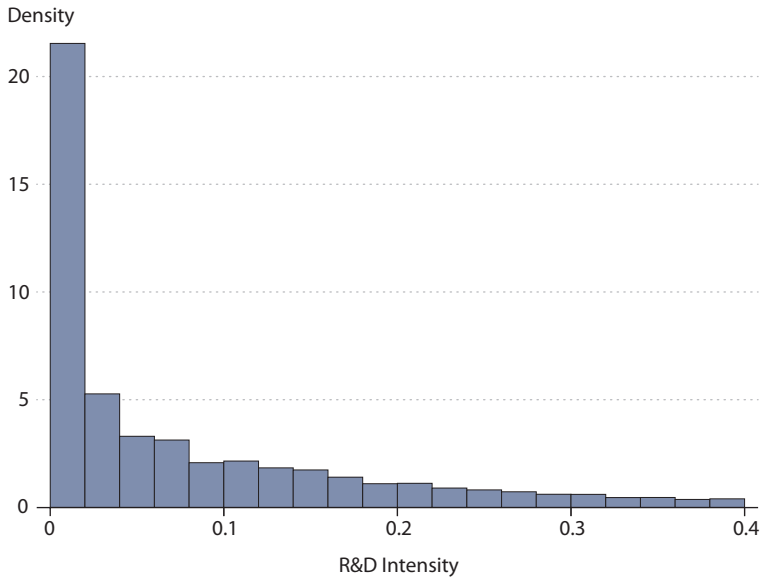
In this section, we explore the importance of several factors in accounting for the cash-holding patterns in the cross section of U.S. firms. For each of these factors, we construct a variable that captures its importance at the firm level. For instance, to evaluate the role of uncertainty, we construct the SD of the growth of total factor productivity at the firm level. Call this variable the proxy. Then we look at the correlation of the proxy variable with cash holdings. We do this at both the firm level and aggregating across sectors. We consider a positive correlation an indication that the factor represented by the proxy variable is important in accounting for cross-sectional variations in cash holdings. Finally, to determine whether this factor may be important in accounting for the trend in cash holdings, we present the evolution over the past 20 years of the proxy variable together with the trend in cash holdings. If the factor under consideration is important in accounting for the recent rise in cash holdings, we would expect the trend of the proxy variable to be similar to that of cash holdings.[2]

Notice that these are correlations, over time and in the cross section of firms, between two endogenous variables. Of course, the interpretation cannot be that the factors we consider *caused* cash holdings to increase. We use the correlations here to try to identify which groups in the population hold more cash and at what time. We think this information will be useful, for instance, in future research attempting to explain these trends with structural models.

We pursue the effects of the following factors on cash holdings: a firm's intensity in research and development (R&D) activities, weight of foreign operations, rate of organizational expenditure, and volatility in firm activities. We also briefly discuss the literature proposing the importance of each of these factors.

Figure 5

Distribution of R&D Intensity (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

Research and Development Intensity

R&D activities are strongly related to the cash holdings of firms because of the inherent uncertain nature of R&D. Availability of projects does not follow a deterministic pattern and good investment opportunities can occur when firms' liquidity is tight. Hence, firms heavily involved in R&D activities are likely to hold more cash than those that are less involved. In fact, Opler et al. (1999) argue that firms with stronger growth opportunities hold more cash and the cash ratio increases with their R&D intensity. Moreover, in a recent article, Pinkowitz, Stulz, and Williamson (2012) confirm the same logic. They show that the recent abnormal cash holdings in the United States—the portion that cannot be predicted by the behavior in the late 1990s—can largely be attributed to firms highly engaged in R&D activities.

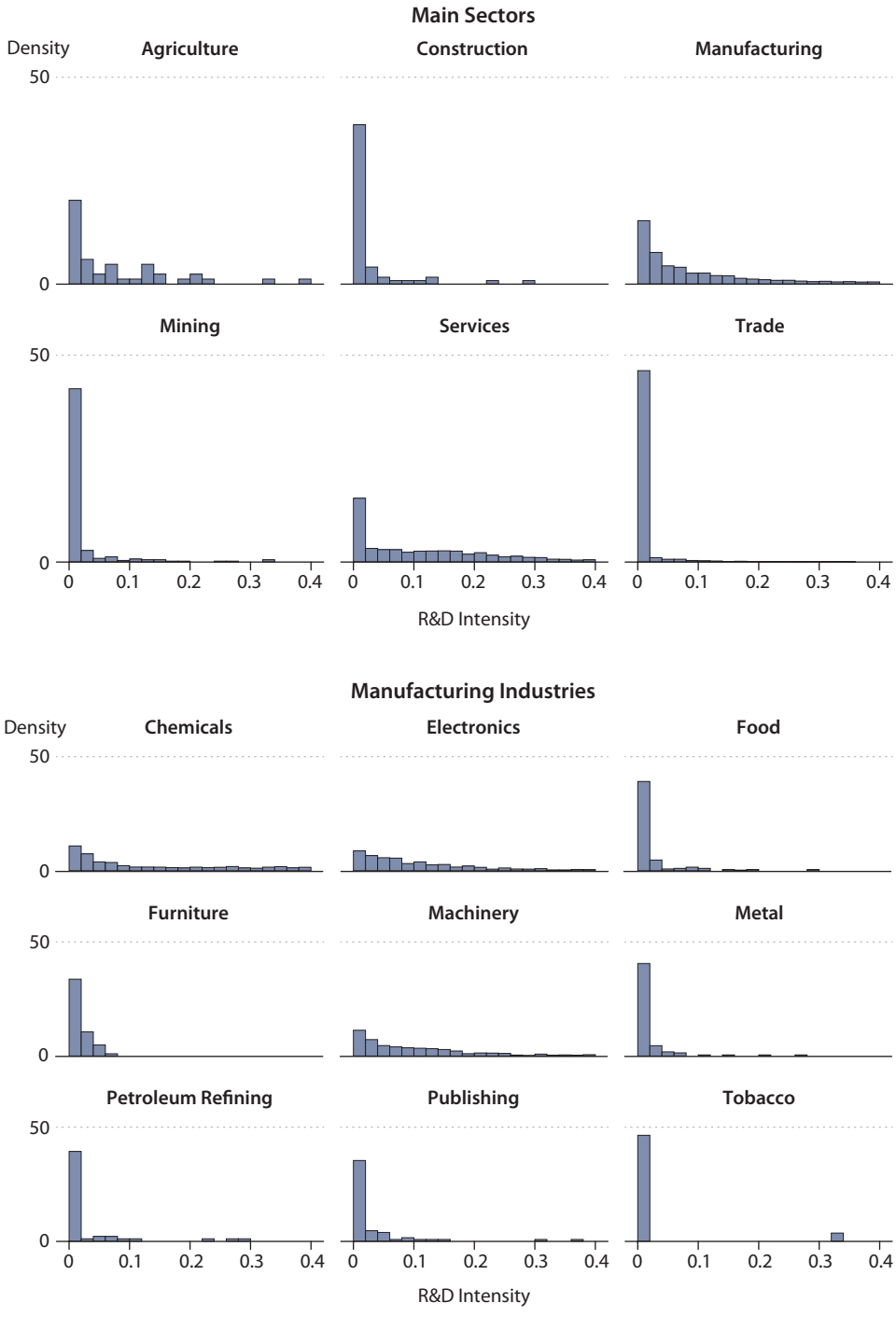
In this section, we explore the role of R&D expenditures in the cash holdings of firms in our sample. We use the R&D expenditures of firms relative to their size to gauge their intensity in research. In other words, the variable we construct to study how heavily firms invest in R&D is

$$R\&DIntensity_{it} = \frac{R\&DExpenditures_{it}}{Assets_{it}}.$$

Figure 5 shows the distribution of R&D intensity in our sample. The distribution is quite skewed: Only 33 percent of the firms have R&D intensity larger than 0.10, the mean ratio in the sample. The top panel of Figure 6 shows the distribution for six specific sectors (agriculture,

Figure 6

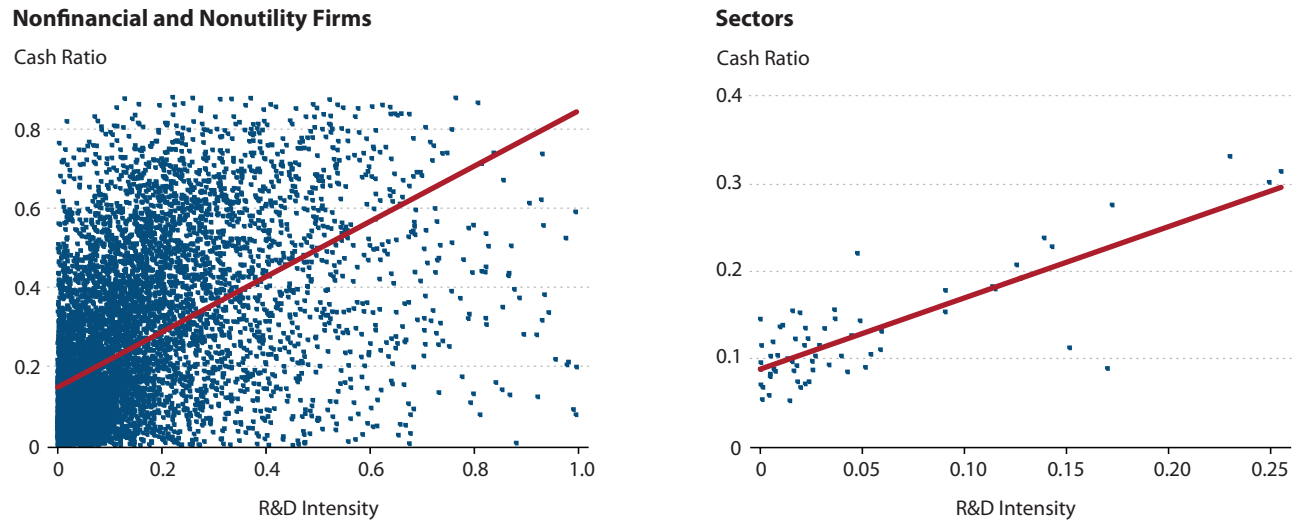
Distribution of R&D Intensity by Sectors



SOURCE: Compustat.

Figure 7

R&D Intensity and Cash Ratio



SOURCE: Compustat.

construction, manufacturing, mining, services, and trade). In most sectors, as in the overall sample, many firms have no R&D expenditures. In some specific sectors, such as manufacturing and services, the distribution is more even—that is, in these sectors many firms are heavily involved in R&D. The SDs of R&D intensity for manufacturing and services are identical, 0.16, while the SD for the whole sample is 0.15. This trend is also reflected in the average R&D expenditures-to-assets ratio. Among the six sectors, services has the highest mean, with 0.14, and manufacturing has the second-highest, with 0.12. The third sector (agriculture) has a much smaller share, with 0.08. As shown in the bottom panel of Figure 6, the manufacturing industries with widely spread distributions for this ratio are chemicals, electronics, and machinery, with average ratios of 0.23, 0.13, and 0.12 and SDs of 0.22, 0.13, and 0.13, respectively.

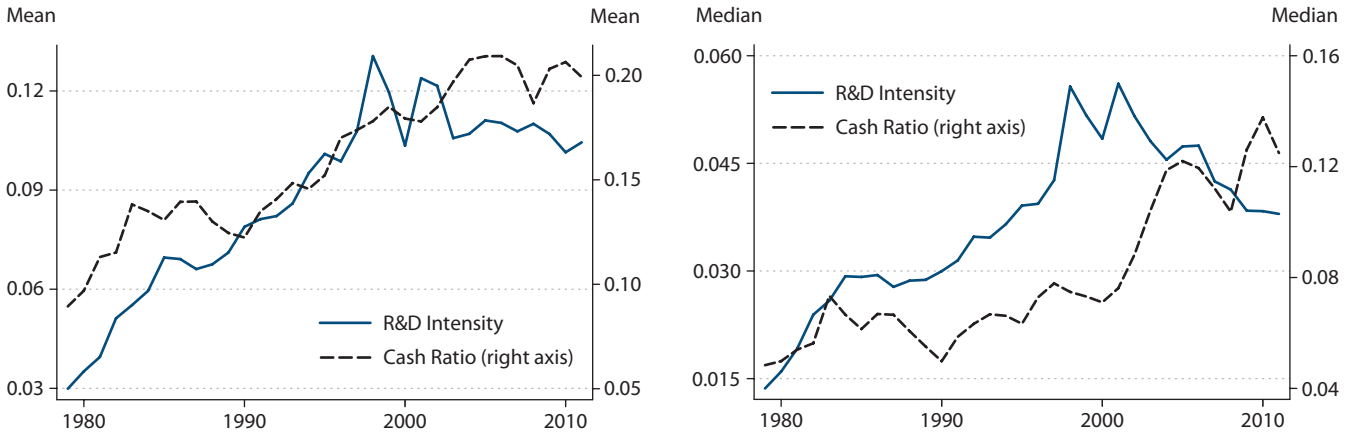
Recall that the sectors with the highest average cash ratio are services and manufacturing; the corresponding industries within manufacturing are chemicals, electronics, and machinery. These are also the industries invested more heavily in R&D.

Figure 7 provides a second demonstration of the association between cash and R&D activities. The left panel plots the mean cash ratio for each firm (across time) against mean R&D intensity. The obvious positive correlation between the two variables holds if firms are collapsed into their two-digit Standard Industrial Classification (SIC) sectors (see the right panel of Figure 7). One two-digit sector in the upper-right corner of the graph with R&D intensities higher than 0.2 and a cash ratio higher than 0.3 is chemicals (in the manufacturing sector), and the other two are membership organizations and miscellaneous services (in the services sector).

Since our main goal is to identify factors that could account for the increase in cash holdings, it is useful to consider the time series of R&D expenditures. Figure 8 shows the behavior of

Figure 8

R&D Intensity (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

R&D expenditures compared with the cash ratio; means are shown in the left panel and medians in the right panel.[3] The figure shows that R&D intensity did indeed increase from the 1980s until the late 1990s. In particular, the annual growth rate of R&D intensity from 1979 to 1998 was around 8 percent for both the mean and the median, which is consistent with the trend in the cash ratio for this interval. Conversely, Figure 8 shows the mean and median R&D intensity for the sample was stagnant over the past decade rather than exhibiting the increasing pattern of the cash ratio. In fact, from 1998 to 2011 the annual rate of decline in the mean R&D intensity is 2 percent; for the median, this rate is 3 percent. If the flow of R&D expenditures is the basis for the recent increase in the stock of cash holdings, we would expect to see increased firm R&D intensity over the same period.

The main point we hoped to convey in this section is that R&D seems important for cash holdings in the cross section of firms. On the other hand, the decreasing trend for this variable over the past decade does not match well with the increase in the cash ratio. This mismatch suggests that although R&D intensity is closely related to the cash ratio in the cross section of firms, it cannot account for the recent trend in cash holdings.

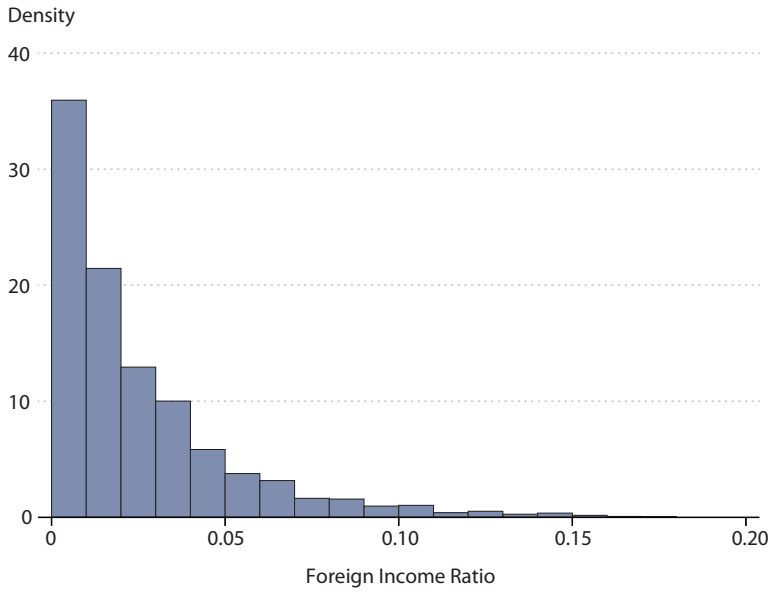
Foreign Income Ratio

Firms that operate heavily abroad may have an additional motivation for cash holding: the tax burden associated with repatriating foreign earnings. In particular, Foley et al. (2007) show this association empirically and argue that firms that would incur more costs by repatriating foreign earnings hold more cash.

Following a similar logic, we study the relationship between the weight of foreign operations in firms' activities and their cash-holding behavior. The variable of focus here is the ratio of pretax foreign income of firms to their total assets:

Figure 9

Distribution of Foreign Income Ratio (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

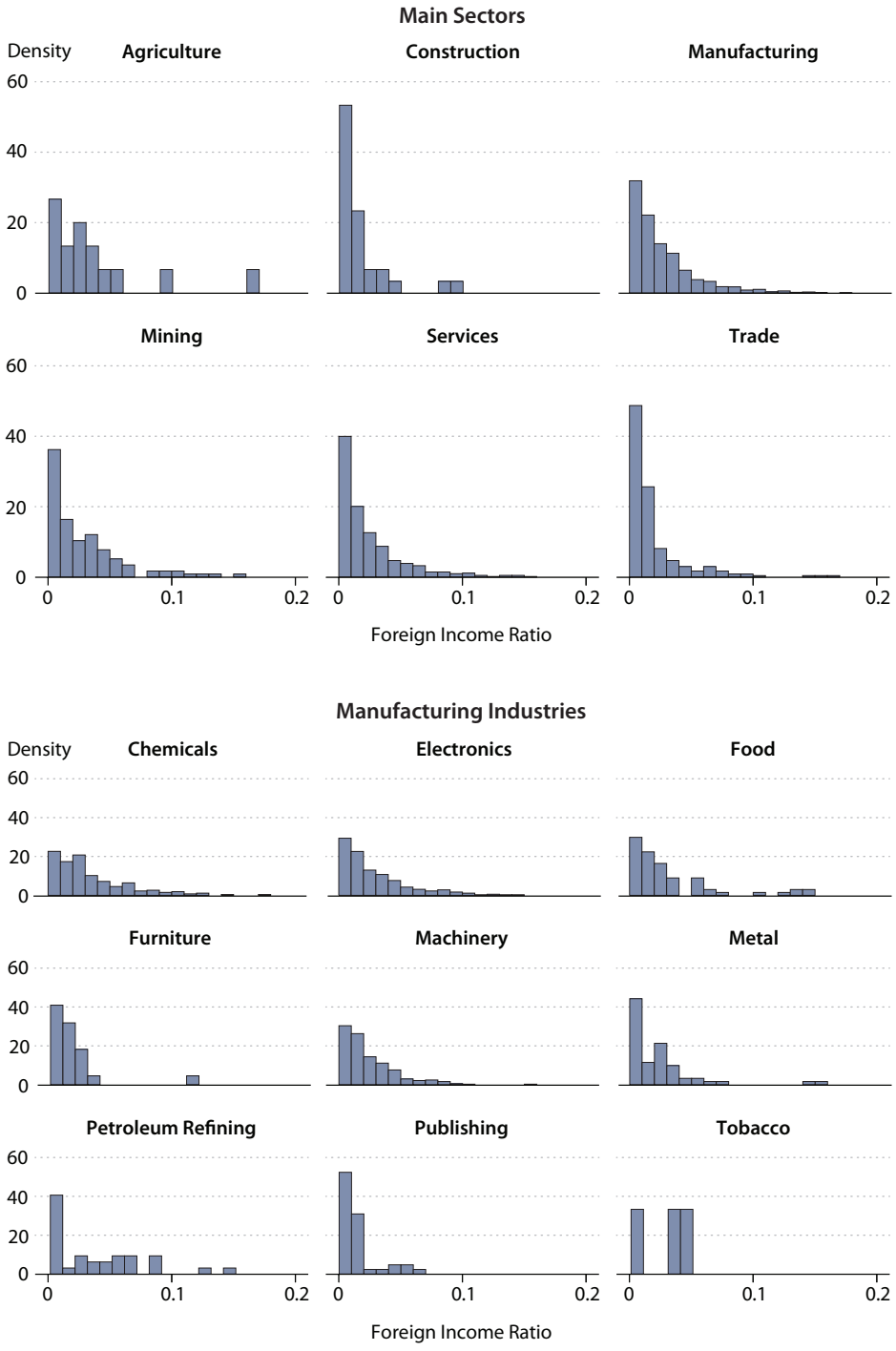
$$ForeignIncomeRatio_{it} = \frac{ForeignIncome_{it}}{Assets_{it}}.$$

As with the R&D intensity, we start by showing the overall distribution of this variable (Figure 9) and do the same for different sectors in the sample (Figure 10, top panel) and industries within manufacturing (Figure 10, bottom panel). As before, there is a concentration near zero in the overall sample. In fact, only 43 percent of the firms have a foreign income ratio larger than 0.01, which is the mean in the sample. In terms of the increased density of the distribution near zero, the heterogeneity across sectors and industries that appears for R&D intensity is absent for the intensity of foreign operations. The mean foreign income ratio for all sectors is between 0.02 (construction, services, and trade) and 0.04 (agriculture), and the sector-specific SD is also between 0.02 and 0.04. The range of means within the manufacturing sector is limited between 0.01 (publishing) and 0.04 (petroleum refining).

The left panel of Figure 11 shows the relation between the cash ratio and foreign activity across firms. If firms do not repatriate cash earned overseas, we would expect to see a strong positive correlation between the cash holdings ratio and foreign income ratio. There is a positive correlation, but it is not as strong as for the R&D intensity. The corresponding graph for two-digit sectors illustrates the same relation (Figure 11, right panel). Although the correlation in this grouping seems more significant, it is mostly due to the vaguely defined two-digit sector of miscellaneous services.

Figure 10

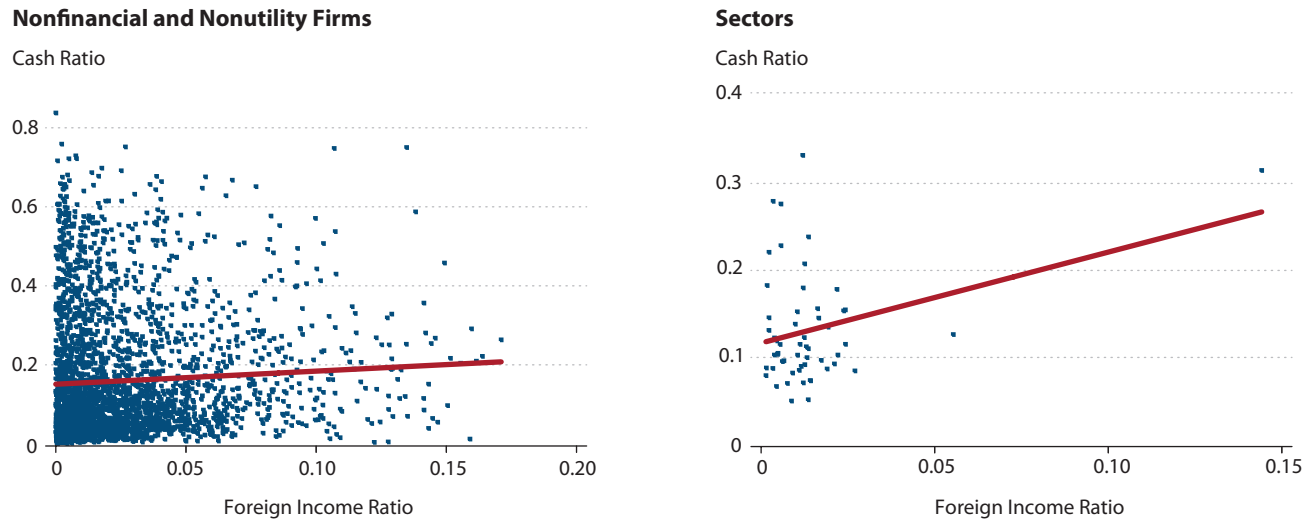
Distribution of Foreign Income Ratio by Sectors



SOURCE: Compustat.

Figure 11

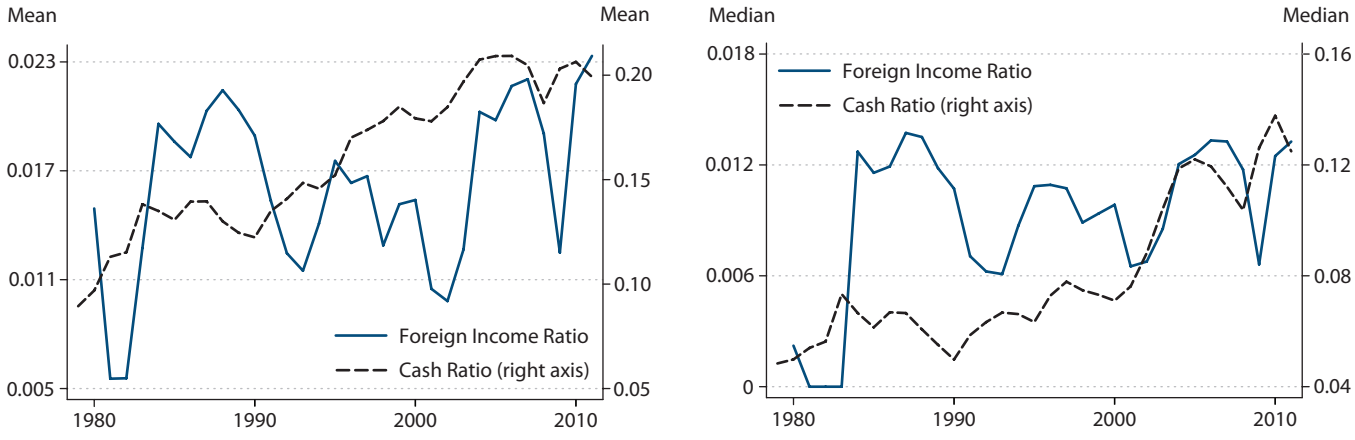
Foreign Income Ratio and Cash Ratio



SOURCE: Compustat.

Figure 12

Foreign Income Ratio (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

Figure 12 illustrates the time series of the foreign income ratio; the left panel corresponds to the mean and the right panel to the median foreign income ratio. If firms do not repatriate cash earned overseas, the cash ratio should increase faster when the foreign income ratio is higher. The evolution of these statistics does not exhibit the clearly increasing behavior of the cash ratio. For instance, from 1990 to 2000, the mean foreign income ratio lacks a persistent slope, while the cash ratio has a clearly positive trend (Figure 12, left panel).

In short, although certain similarities exist between the behavior of the foreign income ratio and cash holdings for the past 10 years, the foreign income ratio in the late 1980s is very similar to that in the 2000s. In addition, the correlation in the cross section of firms is weak, indicating that firms with higher cash ratios are not necessarily those that are more active in foreign operations. Thus, it seems difficult to conclude that the recent rise in cash holdings is an outcome of more firms heavily involved in activities abroad.

Organizational Expenditure Rate

The efficiency of a firm in using its resources (such as human and physical capital) and its capability for future growth depend highly on its current knowledge base and organizational efficiency. Such stock is what we consider as a firm's organizational capital, which scholars have defined in various ways.[4]

Since organizational capital is closely connected to the future growth and investment opportunities of a firm, it may be naturally linked to cash holdings. In fact, Speckbacher and Wentges (2009) argue that "from the management's point of view, high corporate cash holdings allow investments in organizational capital even in periods with poor operating cash flows" (p. 55).

We construct a measure of the organizational expenditure rate as follows[5]:

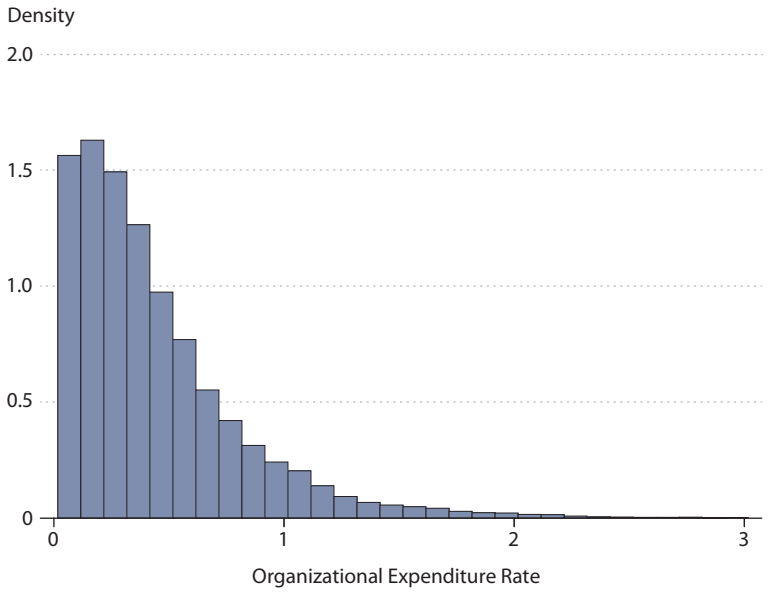
$$OrganizationalExpenditureRate_{it} = \frac{OrganizationalExpenditures_{it}}{Assets_{it}}.$$

Figures 13 and 14 illustrate the distribution of this statistic in the sample. The overall distribution is concentrated in values close to zero. Across sectors, there is considerable heterogeneity; services, trade, and manufacturing have much more widely spread distributions than the other sectors. The SDs within these sectors are 0.43, 0.35, and 0.35, respectively. The mean organization expenditure rates for services, trade, and manufacturing, the highest three in the list, are 0.55, 0.51, and 0.43, respectively. Within manufacturing, the chemicals industry has the highest mean of 0.54 and the flattest distribution. Notice that these groups of firms roughly coincide with those with the highest cash ratios.

Figure 15 more clearly illustrates the positive correlation between the organizational expenditure rate and cash holdings. One example is the outlier two-digit industry membership organizations (shown in the right panel), with an organizational expenditure rate of 0.94 and an average cash ratio of 0.30. The path of this variable across time is qualitatively similar to the R&D intensity. The left panel of Figure 16 shows the evolution of the mean, and the right panel shows that of the median of this ratio compared with the pattern for the cash ratio. A sharp increase occurs in the organizational expenditure rate from the beginning of the sample to 2001. The yearly growth rate from 1979 to 2001 is around 1 percent for both the mean and the median. Never-

Figure 13

Distribution of Organizational Expenditure Rate (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

theless, during the 2001-11 period, organizational expenditure declines at an annual rate of around 3 percent for both statistics.

Analysis of the cross section of firms shows a positive correlation between the organizational expenditure rate and the cash ratio. However, the organizational expenditure rate does not have an increasing pattern comparable to that of the cash ratio for the past decade. Hence, it seems unlikely that the organizational expenditure rate is a major factor in accounting for the rise in cash holdings over the past decade.

Volatility

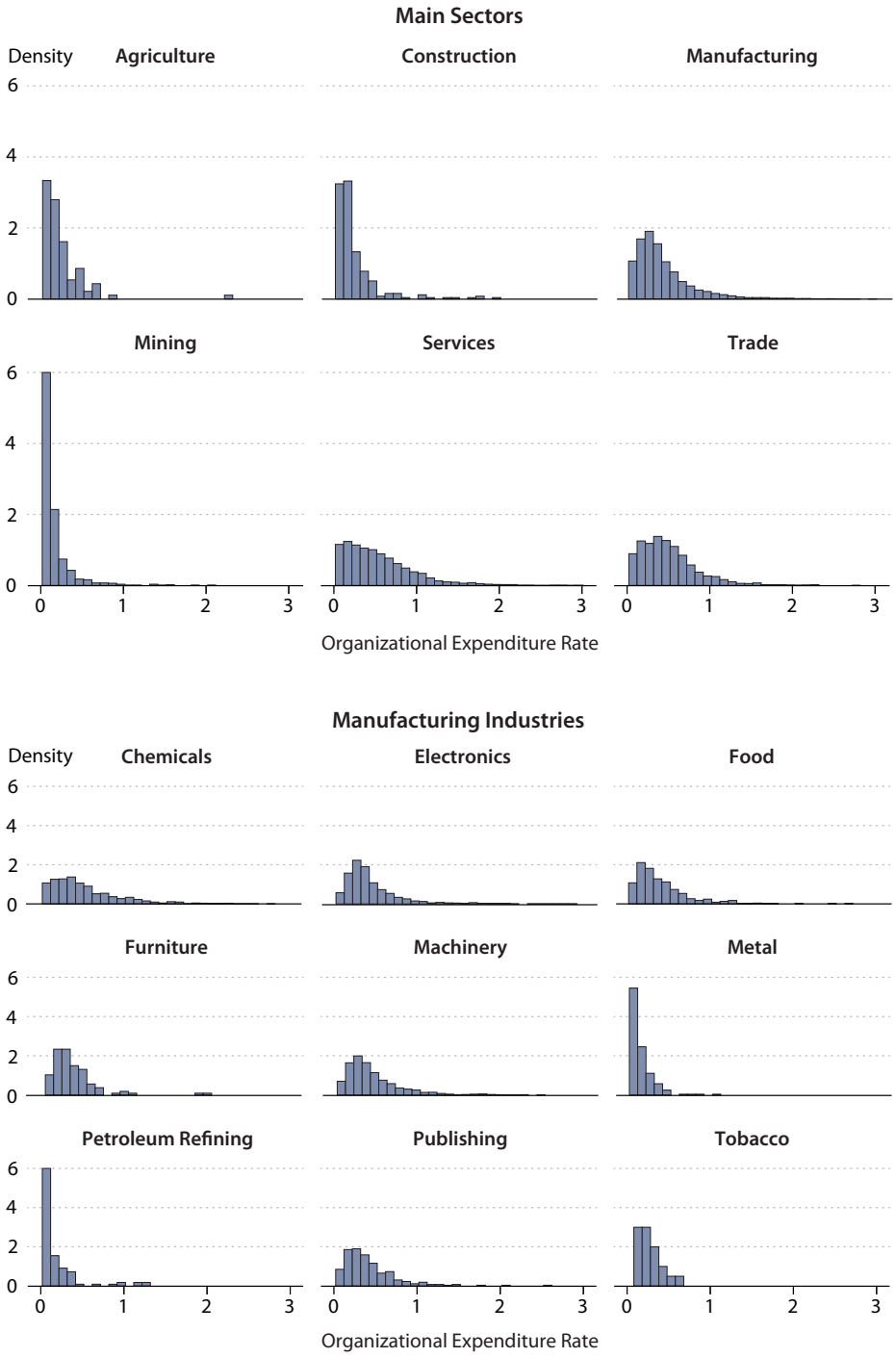
Uncertainty associated with firms' activities may require them to hoard cash for various reasons. Not only will a bad productivity shock affect a firm's cash flow, implying the need for liquid assets, but it is also likely that good shocks boosting the firm's productivity or sales necessitate investment, which requires resources to undertake.

Bates, Kahle, and Stulz (2009) show that the groups of firms with major increases in cash ratios are in industries with the highest increases in idiosyncratic volatility. They conclude that the precautionary motive is important in accounting for higher cash holdings.

In this section, we examine how instability and unpredictability at the individual variable level are related to firms' cash holdings. In particular, we focus on firms' productivity, sales, and cash flow. We construct a measure of volatility for each variable using the variation in the corresponding growth rates.

Figure 14

Distribution of Organizational Expenditure Rate by Sectors



SOURCE: Compustat.

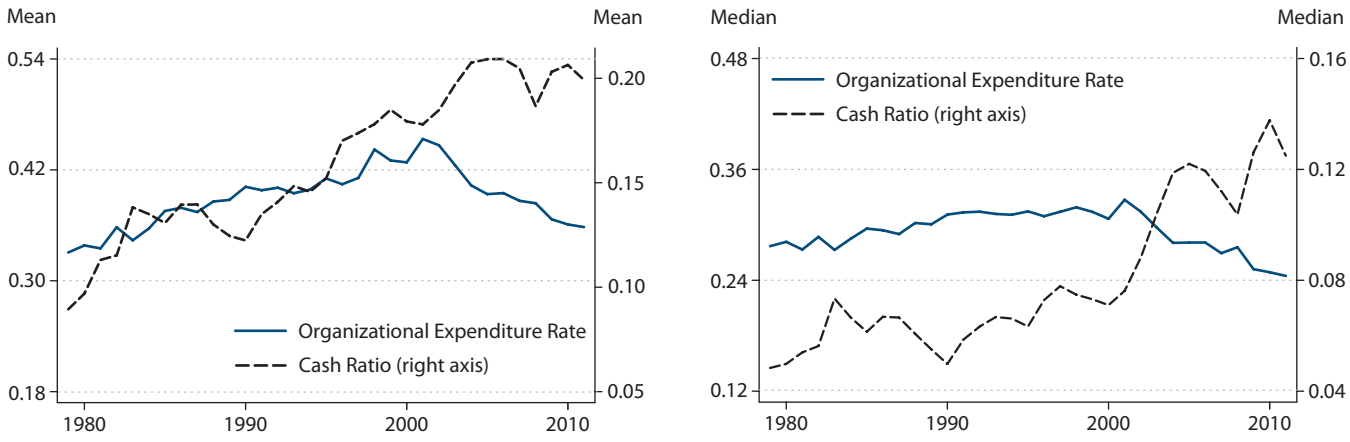
Figure 15

Organizational Expenditure Rate and Cash Ratio

SOURCE: Compustat.

Figure 16

Organizational Expenditure Rate (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

While all three variables used to measure volatility are closely related, it is useful to present different measures to check the robustness of our results. For all analyses, we use growth rates rather than levels to eliminate the effects of trends. Continuation of a trend (increasing or decreasing) is likely to be anticipated and hence less likely to motivate cash holdings. More relevant in this context is the extent of firm-specific variation in the pattern of growth rates for these variables.

In constructing the proxy variables for this section, we assume that firms learn from past experiences. They think that large variations in growth of productivity, sales, and cash flows are possible, and they save cash for that eventuality only if they have experienced such variations in the recent past.

Productivity Volatility. One variable potentially correlated with cash holding is firm-level productivity volatility. We capture productivity by using the residual in the linear regression of the logarithm of operating income on the logarithm of the firm's capital stock:

$$\log\left(OperatingIncome\right)_{it} = \beta_0 + \beta_1 \log\left(Capital\right)_{it} + \varepsilon_{it}.$$

The assumption that firms know this relationship is our reason for running this regression and using variations in the error term to represent uncertainty. Thus, deviations from this relationship are good proxies for surprises. Correspondingly, we define the productivity growth as the change in the residuals of this estimation from the previous year. In turn, the firm-specific volatility in productivity is the SD of this growth rate across the past nine observations of a firm[6]; that is,

$$ProductivityVolatility_{it} = \frac{1}{8}\sum_{j=t-8}^{t}\left(g_{ij} - \overline{g}_i^{(t)}\right)^2,$$

where

$$\overline{g}_i^{(t)} = \frac{1}{9}\sum_{j=t-8}^{t} g_{ij},$$

$$g_{it} = \varepsilon_{it} - \varepsilon_{it-1},$$

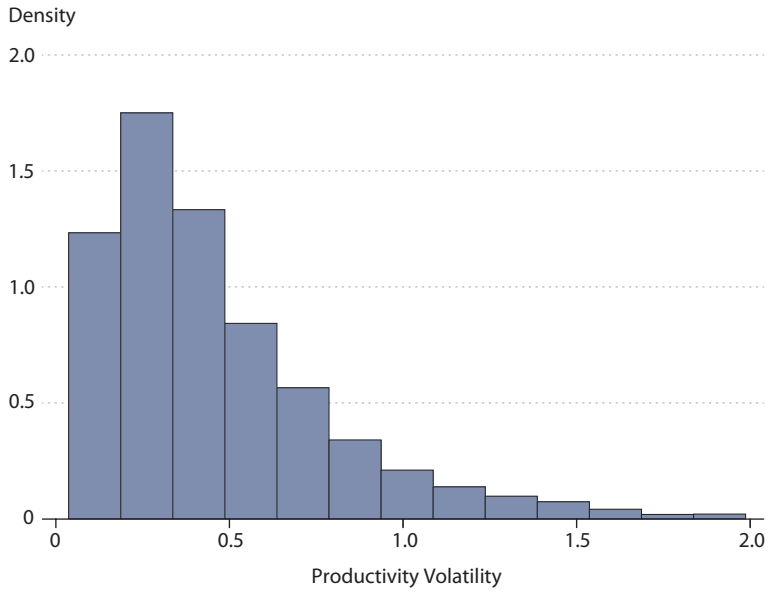
and ε_{it} is the residual in the previous regression.

Figures 17 and 18 illustrate the productivity volatility distribution in the sample. Most sectors show high concentrations of low volatility accompanied by a fat right tail. For the overall sample, 63 percent of the firms have volatility smaller than the mean of 0.47. At the same time, volatility for some sectors and industries is less concentrated close to the zero point. Across sectors, the volatility distribution for construction is the most evenly spread, with a mean of 0.65; the next-closest sector is mining, with an average volatility of 0.54. Within manufacturing, with an average volatility of 0.49, the metal industry has the highest overall volatility with a mean of 0.62.

As with R&D or foreign operations intensity, volatility in productivity is positively correlated with cash holdings in the cross section of firms. However, the aggregate data over two-digit sectors shows a smaller correlation (Figure 19). Industries such as legal services (with a cash ratio more than 0.27 and productivity volatility of 0.2) and building construction general contractors and operative builders (with a cash ratio lower than 0.1 and productivity volatility as high as 0.7) keep the sector-wise correlation from being more positive.

Figure 17

Distribution of Productivity Volatility (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

Figure 20 shows the time series of this productivity volatility.[7] With this simple measure of volatility, the absence of a long-lasting increasing trend in the study period is noticeable. If anything, there is a decrease with an annual rate near 1 percent for both means and medians.

In a nutshell, this measure of volatility exhibits strong positive correlation with the cash ratio in the cross section of firms, while an increasing recent trend to accompany the recent pattern in cash holdings is missing. Thus, it seems difficult to argue that idiosyncratic risk in productivity can account for the recent trend in cash holdings.

Sales Volatility. In this section, we construct a measure of firm-specific volatility using sales. In particular, we focus on the SD of the change in the logarithm of sales within the past nine observations for a firm:

$$SalesVolatility_{it} = \frac{1}{8}\sum_{j=t-8}^{t}\left(d_{ij} - \overline{d}_i^{(t)}\right)^2,$$

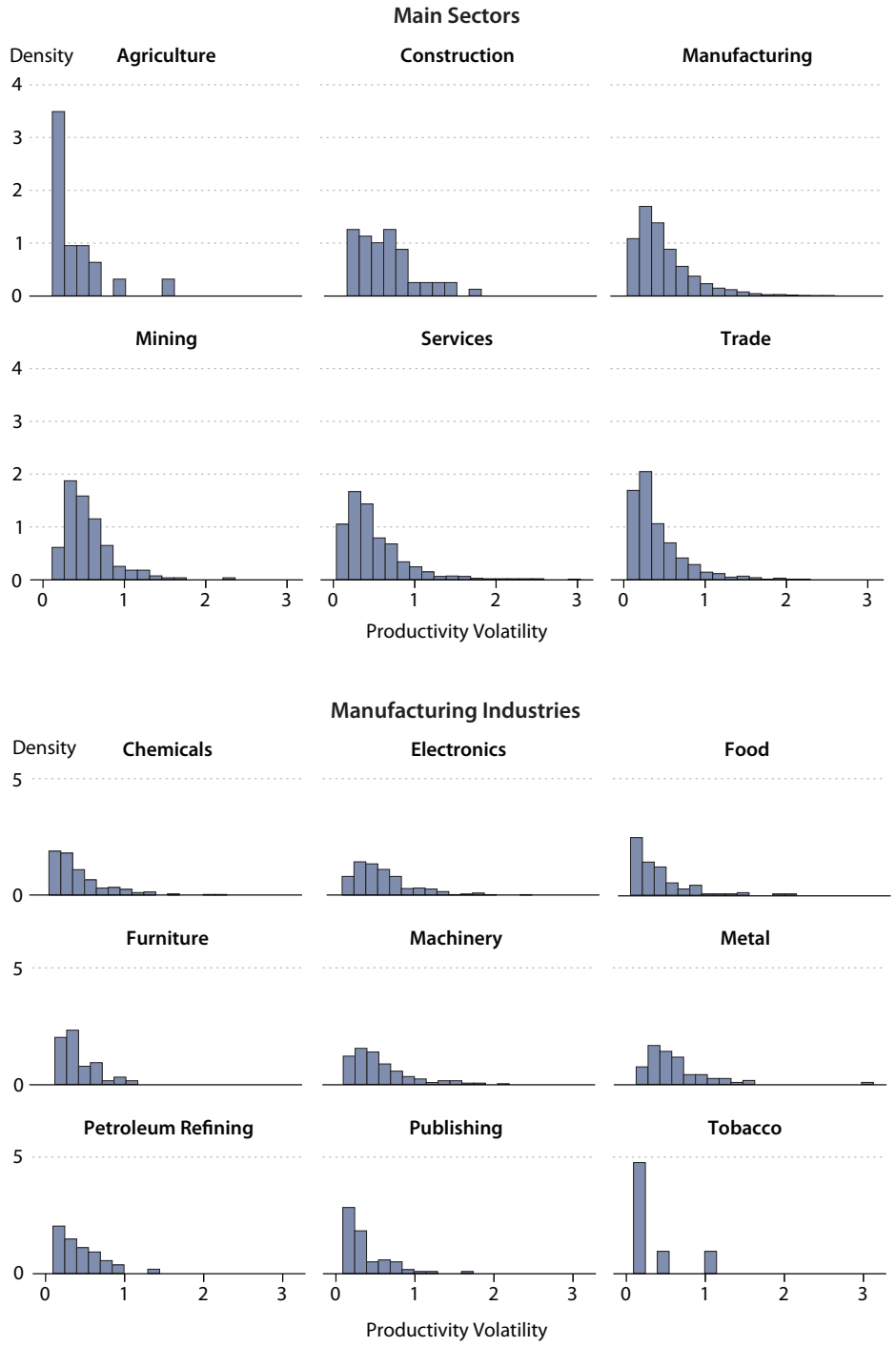
where

$$\overline{d}_i^{(t)} = \frac{1}{9}\sum_{j=t-8}^{t} d_{ij}, \text{ and}$$

$$d_{it} = \log(Sales)_{it} - \log(Sales)_{it-1}.$$

Figure 18

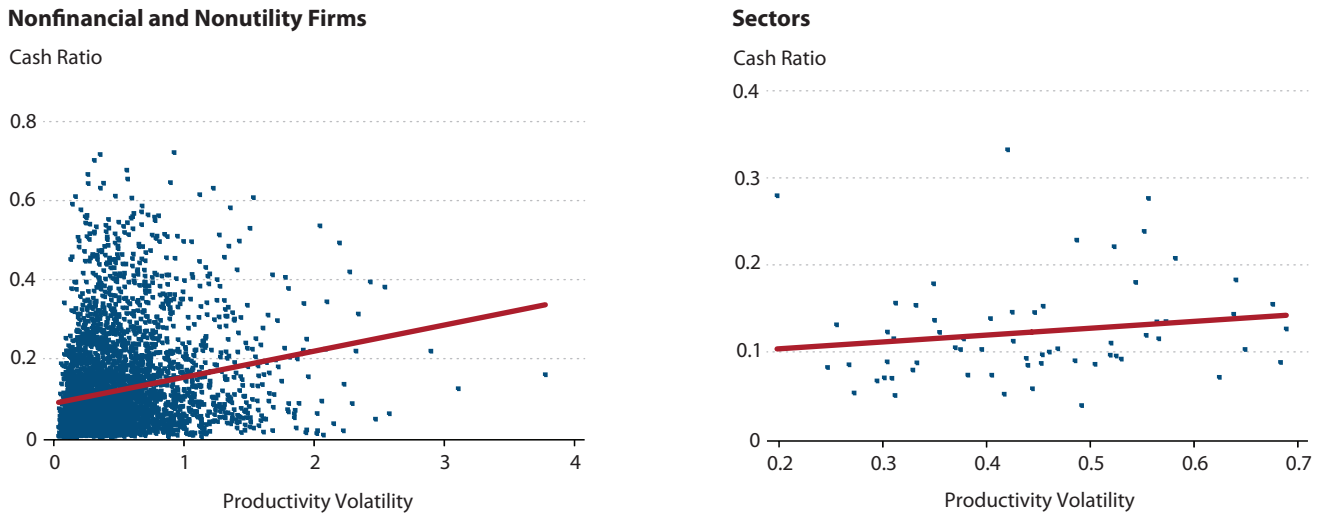
Distribution of Productivity Volatility Rate by Sectors



SOURCE: Compustat.

Figure 19

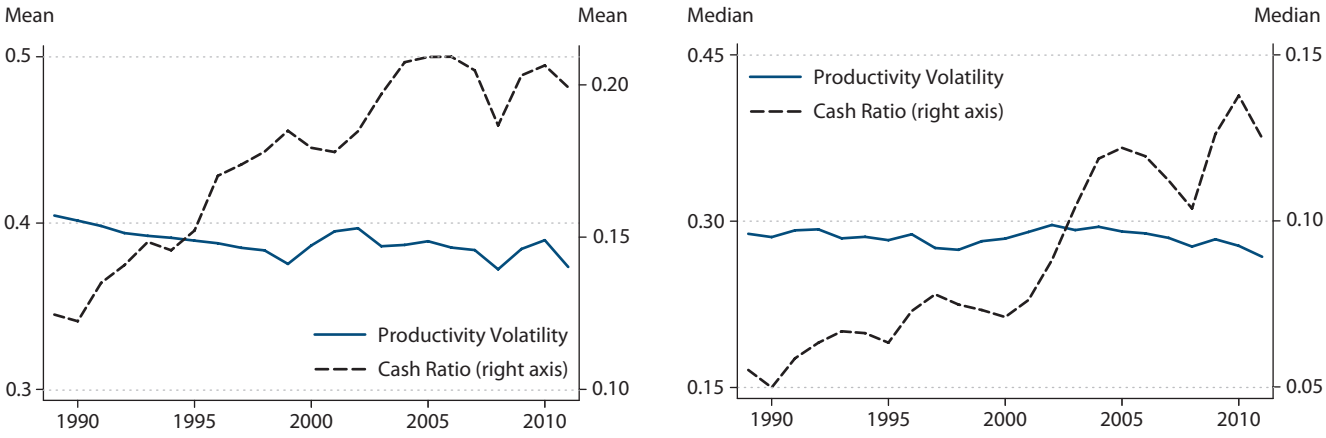
Productivity Volatility and Cash Ratio



SOURCE: Compustat.

Figure 20

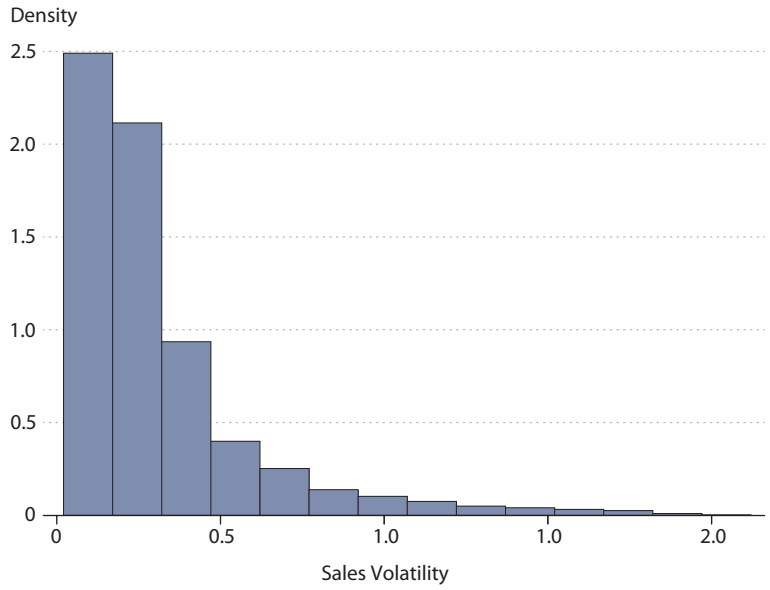
Productivity Volatility (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

Figure 21

Distribution of Sales Volatility (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

Figure 21 shows the distribution of sales volatility throughout the firms in the sample. In addition to the fat right tail, the similarity of this distribution to the distribution of productivity volatility is noticeable. For this measure, volatility is lower than the mean of 0.33 for 69 percent of firms. According to this measure, the mining sector has the highest volatility, with an average of 0.49, which can be seen by observing the less-concentrated distribution for this group near zero (Figure 22, top panel). Meanwhile, some sectors have higher SDs than mining (0.35) (e.g., manufacturing, 0.37). The bottom panel of Figure 22 shows that within the manufacturing sector, which has a mean ratio of 0.33, the chemicals industry has the highest average volatility, with a mean of 0.53, and the flattest distribution.

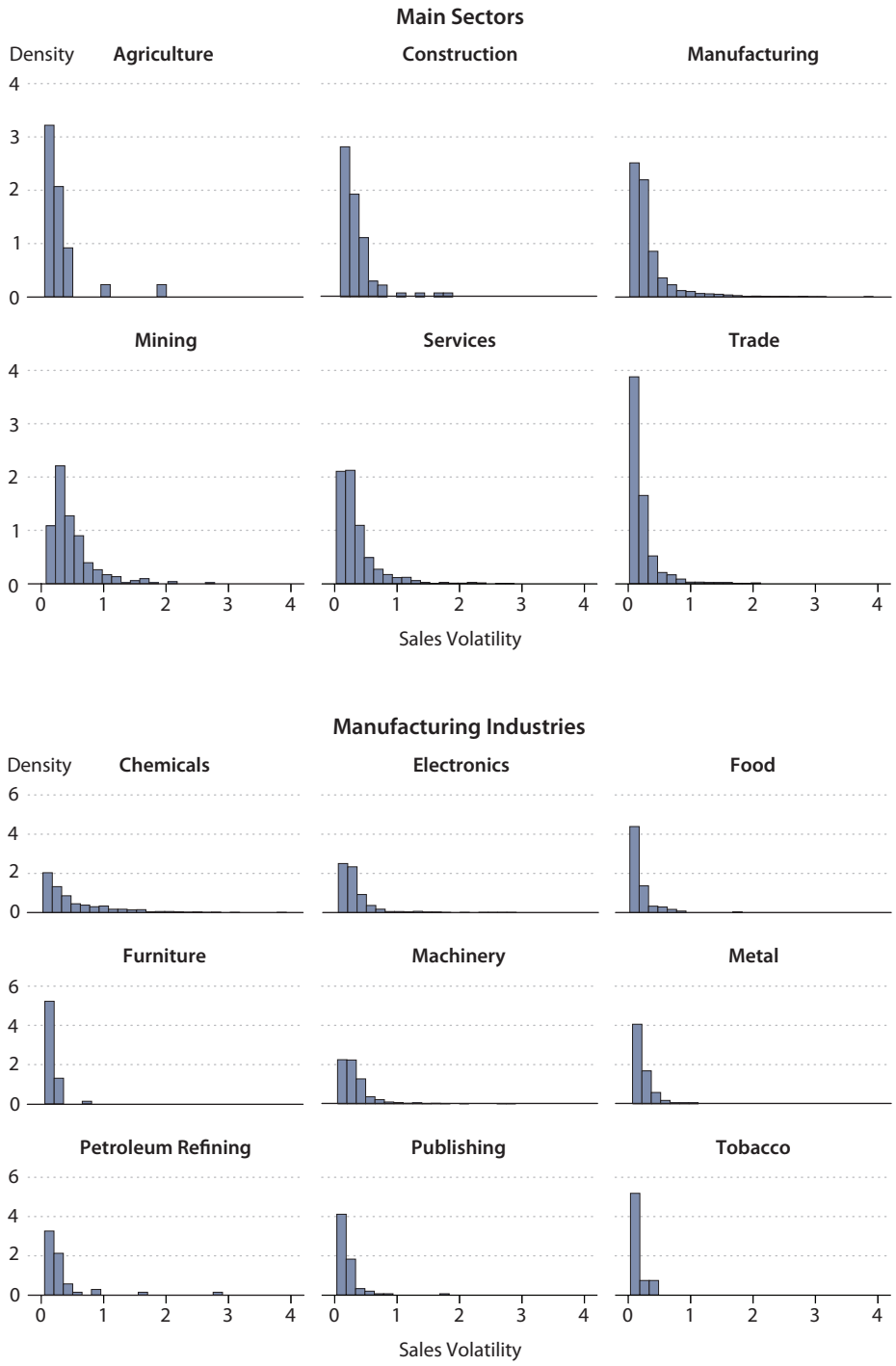
As with the volatility in productivity, there is a positive correlation between sales volatility and cash holding at the firm level (Figure 23, left panel). The right panel shows a positive correlation across two-digit sectors, even though this association appears less significant in the linear fit, partially because of the low cash ratio of forestry (0.08) accompanied by high sales volatility (0.73).

Figure 24 shows the evolution in sales volatility over time, with an increase until 2002, followed by stagnation and then a decline in the past 5 years. This pattern suggests that uncertainty may account for the rise in cash holdings until the early 2000s, but it cannot be associated with the evolution of cash holdings over the past decade.

Cash Flow Volatility. In addition to the indicators previously discussed, another—and perhaps more direct—way to study the relationship between cash holding and volatility in firms'

Figure 22

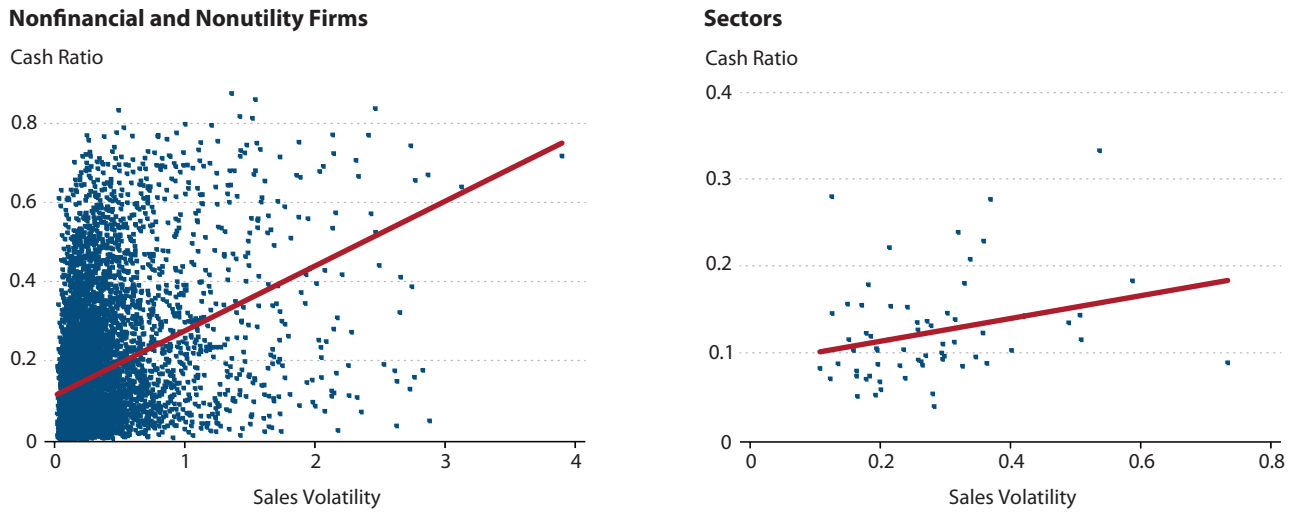
Distribution of Sales Volatility Rate by Sectors



SOURCE: Compustat.

Figure 23

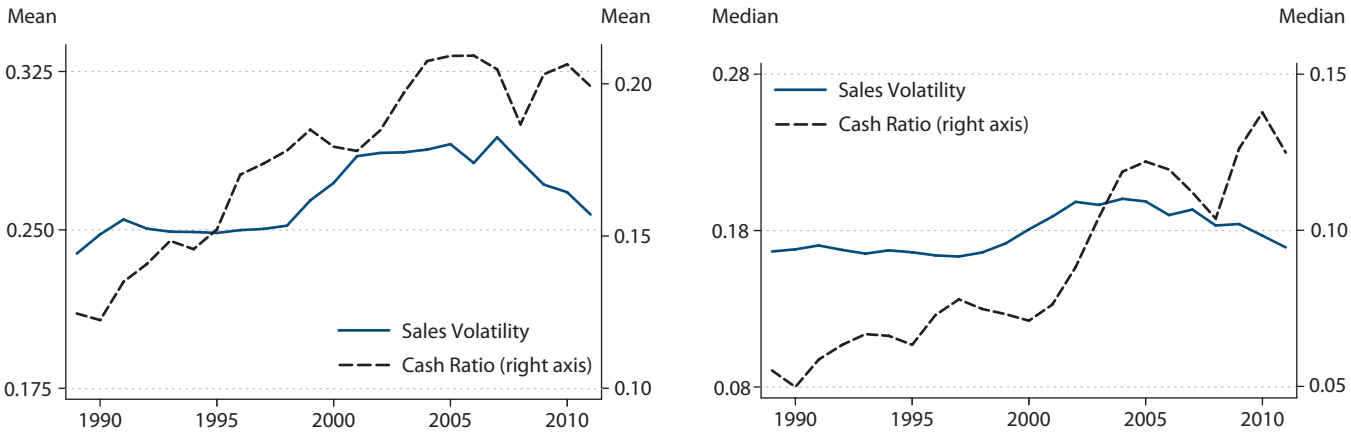
Sales Volatility and Cash Ratio



SOURCE: Compustat.

Figure 24

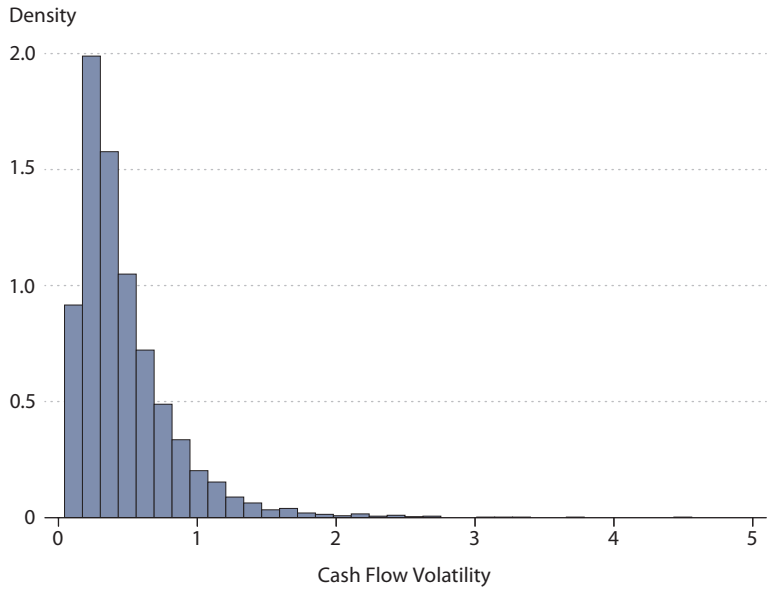
Sales Volatility (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

Figure 25

Distribution of Cash Flow Volatility (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

activities is simply to look at cash flow volatility. Here we define "cash flow" as earnings before interest after deducting income taxes paid and cash flow of dividends. Similar to the analysis presented above, firm-specific volatility of this variable is the SD over time for each firm:

$$CashFlowVolatility_{it} = \frac{1}{8}\sum_{j=t-8}^{t}\left(b_{ij} - \overline{b}_i^{(t)}\right)^2,$$

where

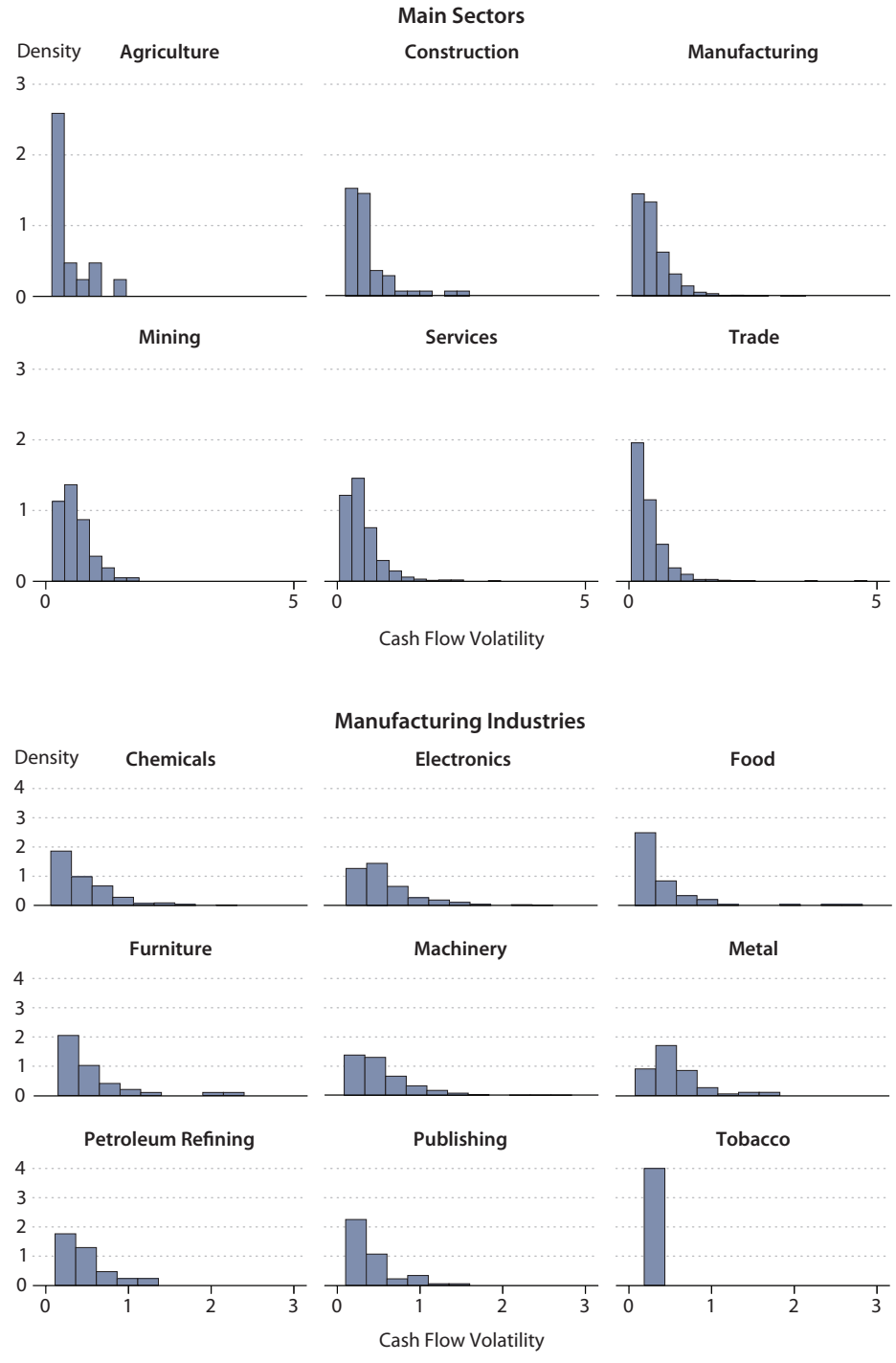
$$\overline{b}_i^{(t)} = \frac{1}{9}\sum_{j=t-8}^{t} b_{ij}, \text{ and}$$

$$b_{it} = \log(CashFlow)_{it} - \log(CashFlow)_{it-1}.$$

Since the measures of volatility capture similar factors, the patterns of distribution for this measure look similar to earlier ones with minor differences. The overall distribution shows very little volatility in the cash flows of many firms (Figure 25). Sector-wise analysis indicates that construction and mining have more firms with high volatility and higher means (0.61 and 0.59, respectively). The highest dispersion among sectors is for mining, with an SD of 0.48, which is much higher than the SD of 0.36 for the entire sample (Figure 26). Within manufacturing, the electronics industry has the flattest distribution and the highest mean, 0.57.

Figure 26

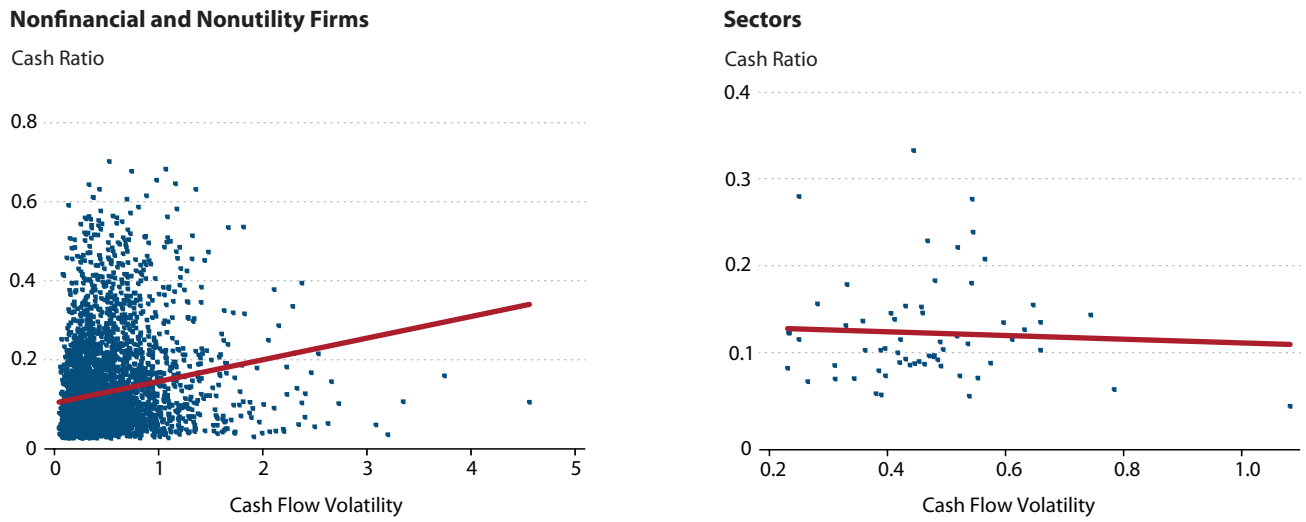
Distribution of Cash Flow Volatility Rate by Sectors



SOURCE: Compustat.

Figure 27

Cash Flow Volatility and Cash Ratio



SOURCE: Compustat.

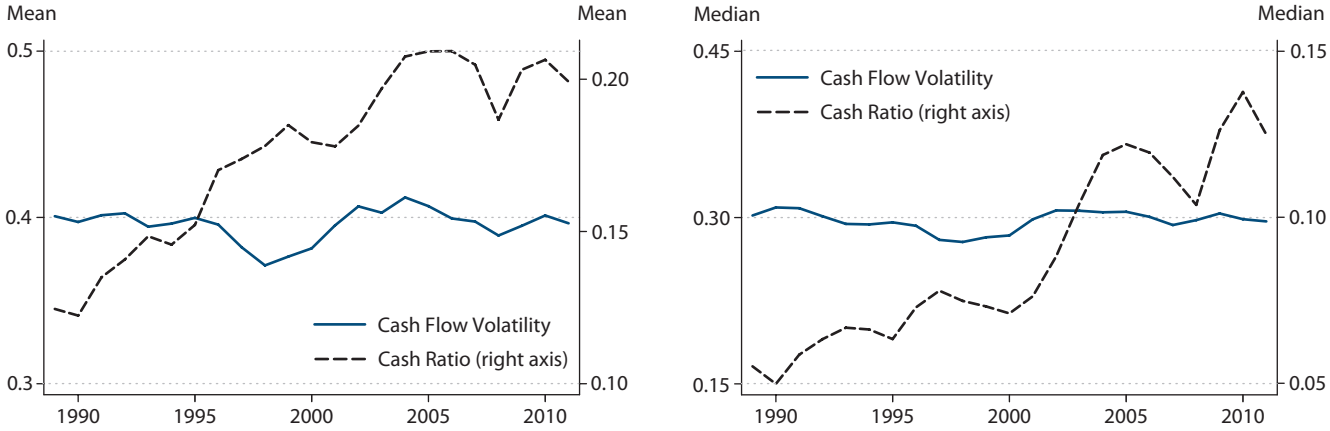
The left panel of Figure 27 shows a positive correlation of cash ratio with cash flow volatility across firms that resembles earlier measures of volatility. The right panel shows that the correlation across two-digit sectors is not positive, as it was for volatilities in productivity or sales. Two broad industries account for this difference: (i) fishing, hunting, and trapping and (ii) local and suburban transit and interurban highway passenger transportation, which have cash ratios around 0.04 and cash flow volatility around 1.

Next we focus on the behavior of this measure of volatility across time. The overall pattern is nothing but steady (Figure 28). With the exception of an interval from 1998 to 2003 characterized by an annual growth rate of 2 percent in terms of the mean, there is dominant stagnation if volatility is captured by firms' cash flows.

Previously we discussed two different measures of idiosyncratic volatility that are constructed using different underlying sources—namely, productivity and sales. Within these two measures, productivity volatility was an example of a measure without a clearly increasing pattern to accompany the trend in the cash ratio, while sales volatility provided both cross-sectional and intertemporal evidence that higher volatility in firms' operations can partially account for higher cash holdings except in recent years. The measure constructed here using firms' cash flows is closer to the former, in the sense that it shows a positive correlation in the cross section but lacks a persistent increasing pattern throughout the study interval. Thus, the analysis here casts doubts about the role of firm-level volatility in accounting for the recent increase in cash holdings.

Figure 28

Cash Flow Volatility (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

THE ROLE OF AGGREGATE UNCERTAINTY

In the previous section, we focused on idiosyncratic factors that could potentially account for cash hoarding. In a nutshell, we found little evidence to support the concept that such motivations are sufficient to account for the recent rise in cash holdings. In this section, we question the relevance of aggregate uncertainty in explaining the recent cash-hoarding phenomenon. Specifically, we focus on aggregate productivity uncertainty and policy uncertainty to study the potential for these factors to account for the recent rise in the cash ratio.

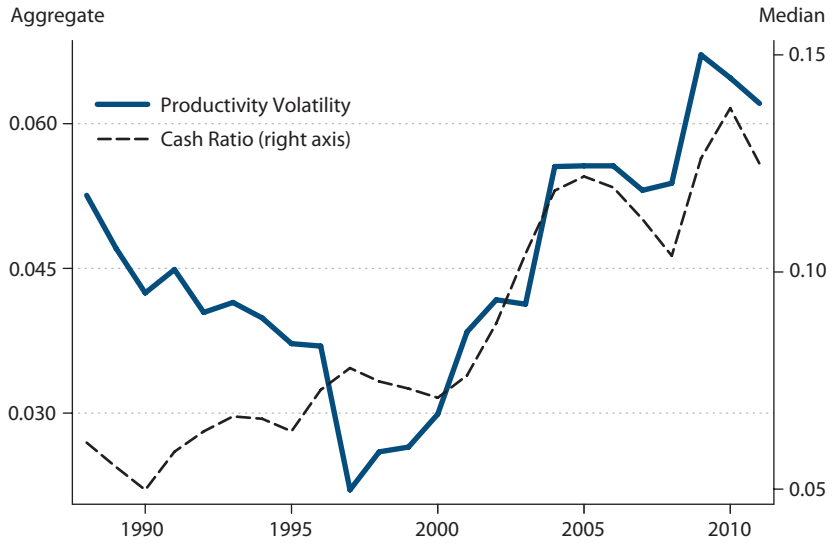
Aggregate Productivity Volatility

One indication of precautionary motives for hoarding cash can be the effects of aggregate volatility in productivity on firms' cash-holding behavior. In an earlier paper, Palazzo (2012) studies the relationship between firms' cash holdings and aggregate risk. He shows that firms with a positive correlation between cash flows and the aggregate shock build cash stocks as a cushion, in line with the previously explained precautionary motives. In this section, we conduct an alternative exercise by comparing the trend in aggregate productivity uncertainty and cash holdings.

Specifically, we construct a measure of volatility in aggregate productivity and show the evolution of this variable over time. In constructing this measure, we use the information on sales and number of employees of each firm in the Compustat sample. For each year we (i) aggregate sales and employment in the sample and (ii) use sales per worker as a measure of productivity. Then, similar to the volatility measure used previously, we compute the log-deviation from the previous year and calculate the measure of volatility as the SD in the past 9 years:

Figure 29

Aggregate Productivity Volatility (Nonfinancial and Nonutility Firms)



SOURCE: Compustat.

$$AggregateProductivityVolatility_{it} = \frac{1}{8}\sum_{j=t-8}^{t}\left(G_{ij} - \overline{G}_i^{(t)}\right)^2,$$

where

$$\overline{G}^{(t)} = \frac{1}{9}\sum_{j=t-8}^{t} G_j$$

$$G_t = \log\left(AggregateProductivity\right)_{it} - \log\left(AggregateProductivity\right)_{it-1},$$

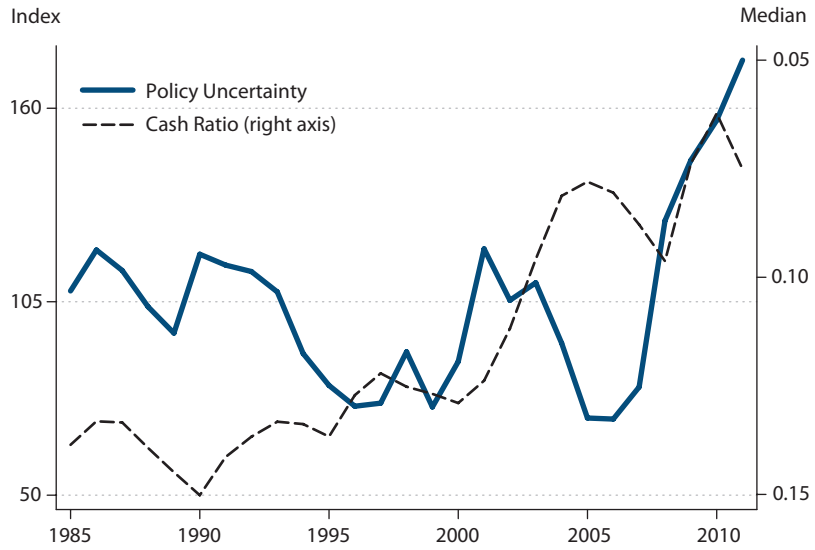
$$AggregateProductivity_{it} = \frac{\sum_{i=1}^{N} Sales_{it}}{\sum_{i=1}^{N} Employees_{it}},$$

and N is the number of firms in the sample.

Notice that here again, as for firm-level volatility, we use past volatility as the measure that firms take into account in their decisions regarding cash holdings. The idea is that firms learn from experience and that firm managers in the early 2000s did not know that the Great Recession was coming.[8]

Figure 30

Aggregate Policy Uncertainty (Nonfinancial and Nonutility Firms)



SOURCE: Baker, Bloom, and Davis (2013).

Figure 29 shows how the evolution of aggregate productivity volatility overlaps with that of the median cash ratio: The movements of the two variables coincide over the past decade. The aggregate productivity volatility in 2011 is 62 percent more than in 2001, while this rate is 64 percent for the median cash ratio. Consequently, the annual growth rates across these 10 years (beginning with 2001) are around 5 percent for both. These two statistics do not correlate well for the previous periods, however.

Two facts point to aggregate productivity volatility as a candidate in accounting for the recent trend in cash holdings: (i) the good correlation between firm-level uncertainty and firm-level cash holdings and (ii) the increasing trends in volatility in aggregate productivity. While these facts are certainly not a proof, they do hint that firms' cash-holding behavior may be affected by uncertainty associated with aggregate productivity.[9]

Aggregate Policy Uncertainty

Next, we use an excellent source to determine how (i) uncertainty associated with policy moves across time and (ii) it relates to the trend in cash holdings. In particular, we use the Economic Policy Uncertainty Index constructed by Baker, Bloom, and Davis (2012). They combine several measures to form this index: the frequency of articles in large newspapers discussing policy-related economic uncertainty, forthcoming expirations of tax code provisions, and disagreement among professional forecasters in predicting some economic statistics.

Figure 30 plots the evolution of the index against time, together with the median cash ratio. Although it is difficult to argue for overall comovement, there are several episodes where the

two variables move in the same direction, in particular between 2008 and 2010. One certain conclusion is that if we compare the correlation with the cash ratio of two measures of aggregate uncertainty we use here with that of the idiosyncratic factors presented previously, aggregate factors seem to dominate.

CONCLUSION

We analyzed data on publicly traded U.S. firms to shed light on the increasing trend in their cash holdings. By 2011, U.S. firms were holding four times as much cash as they were holding in 1995. Perhaps even more striking, the cash-to-asset ratio more than doubled between 2000 and 2010.

We first evaluated the role of idiosyncratic factors. The results suggest that some of these factors—namely, R&D expenditures and cash-flow uncertainty—are likely to explain differences in the cross section and the early part of the aggregate trend. However, these factors stop short of keeping up with the rise in the cash-to-asset ratio that occurred after the mid-2000s. We then focused on the recent trends in aggregate uncertainty, particularly regarding aggregate productivity and policy. We find evidence suggesting the recent increase in cash holdings may be related to these aggregate factors.

DATA APPENDIX

We use Compustat as the data source for our calculations. Only U.S. firms are considered. Unless otherwise noted, we exclude financial firms (SIC classification between 6000 and 6999) and utility firms (SIC classification between 4900 and 4999), as is standard in the literature. Firm-year observations with negative total assets, sales, or capital are dropped. We trim from two tails at the 1 percent level the ratios of the following: cash to total assets, foreign income to total assets, R&D expenditure to total assets, and organizational intensity to total assets.

For cash holdings we use the variable "cash and short-term investments, che," and for assets we use "total assets, at." For R&D expenditures we use "research and development expense, xrd," and for foreign income we use variable "pretax foreign income, pifo." The variable "property, plant and equipment (total (net)), ppent" is used for capital, and "operating income before depreciation, oibdp" is used for operating income. The variable "sales (net), sale" corresponds to the sales of a firm. Organizational capital expenditure is measured by "selling, general and administrative expense, xsga." We use "employees, emp" for a firm's employment. In order to calculate cash flows, we use "earnings before interest, ebitda," "total income taxes, txt," and "cash flow of dividends, dv."

NOTES

1 The appendix describes the data we use and provides the sampling restrictions.

2 The implicit assumption for this logic to follow is that there is a stable relationship between these variables.

3 All figures of intertemporal patterns with two *y*-axes preserve the same range relative to the average value of each particular series to facilitate the comparison.

4 See Black and Lynch (2005) for a survey of the definitions of organizational capital.

5 See Lev and Radhakrishnan (2005) for a similar but more complex measure.

6 In addition to using the past nine observations, we performed similar calculations with the past 5 and 11 observations with no change in results.

7 Since we use the previous eight observations together with the current one as a reference of volatility and then take the difference, the first observation is 1988 instead of 1979 as in the previous variables.

8 This is in line with a large literature in economics in the early 2000s referring to the previous decade as the "Great Moderation."

9 One can note that the evolution of the aggregate volatility is quite different from the behavior of firm-specific volatility. To be clear, neither the fall during the 1990s nor the rise in the 2000s in aggregate productivity volatility is matched by corresponding patterns in firm-level volatility. These findings are in line with those of Comin and Mulani (2004) and Comin and Philippon (2006), who document the discrepancies between the behavior of aggregate and idiosyncratic volatility during the 1990s.

REFERENCES

Baker, Scott R.; Bloom, Nicholas and Davis, Steven J. "Policy Uncertainty: A New Indicator." *CentrePiece—The Magazine for Economic Performance*, No. CEPCP362, Winter 2011/12, *16*(3), pp. 21-23; http://cep.lse.ac.uk/pubs/download/cp362.pdf.

Bates, Thomas W.; Kahle, Kathleen M. and Stulz, René M. "Why Do U.S. Firms Hold So Much More Cash Than They Used To?" *Journal of Finance*, October 2009, *64*(5), pp. 1985-2021.

Black, Sandra E. and Lynch, Lisa M. "Measuring Organizational Capital in the New Economy," in Carol Corrado, John Haltiwanger, and Daniel Sichel, eds., *Measuring Capital in the New Economy* (National Bureau of Economic Research Studies in Income and Wealth). Volume 65, Chap. 6. Chicago: University of Chicago Press, 2005, pp. 205-36.

Brunell, Don C. "Washington View: Uncertainty Keeping Businesses' Cash on the Sidelines." *Columbian* (Vancouver, WA), August 16, 2011; http://www.columbian.com/news/2011/aug/16/uncertainty-keeping-businesses-cash-on-the-sidelin/.

Comin, Diego and Mulani, Sunil. "Diverging Trends in Macro and Micro Volatility: Facts." NBER Working Paper No. 10922, National Bureau of Economic Research, November 2004; http://www.nber.org/papers/w10922.pdf.

Comin, Diego and Philippon, Thomas. "The Rise in Firm-Level Volatility: Causes and Consequences," in Mark Gertler and Kenneth Rogoff, eds. *NBER Macroeconomics Annual 2005*. Volume 20. Cambridge, MA: MIT Press, 2006, pp. 167-228; http://www.nber.org/chapters/c0072.pdf.

Foley, Fritz C.; Hartzell, Jay C.; Titman, Sheridan and Twite, Garry. "Why Do Firms Hold So Much Cash? A Tax-Based Explanation." *Journal of Financial Economics*, December 2007, *86*(3), pp. 579-607.

Lev, Baruch and Radhakrishnan, Suresh. "The Valuation of Organization Capital," in Carol Corrado, John Haltiwanger, and Daniel Sichel, eds., *Measuring Capital in the New Economy* (National Bureau of Economic Research Studies in Income and Wealth). Volume 65, Chap. 3. Chicago: University of Chicago Press, 2005, pp. 73-110.

Opler, Tim; Pinkowitz, Lee; Stulz, René and Williamson, Rohan. "The Determinants and Implications of Corporate Cash Holdings." *Journal of Financial Economics*, April 1999, *52*(1), pp. 3-46.

Palazzo, Berardino. "Cash Holdings, Risk, and Expected Returns." *Journal of Financial Economics*, April 2012, *104*(1), pp. 162-85.

Pinkowitz, Lee; Stulz, René M. and Williamson, Rohan. "Multinationals and the High Cash Holdings Puzzle." NBER Working Paper No. 18120, National Bureau of Economic Research, June 2012; http://www.nber.org/papers/w18120.pdf?new_window=1.

Speckbacher, Gerhard and Wentges, Paul. "A Note on Stakeholder Theory and Risk: Implications for Corporate Cash Holdings and Dividend Policy." *Frontiers in Finance and Economics*, April 2009, *6*(1), pp. 51-72.

Donor Motives for Foreign Aid

Subhayu Bandyopadhyay and E. Katarina Vermann

The literature on foreign aid has contributed to our understanding of the motives for developed nations to provide aid to developing nations. In this article, the authors primarily focus on donor motivation, but they also touch on the consequences of receiving aid for developing nations. They consider both the developmental and strategic aspects of giving aid. While aid in the 1960s focused more on development, recent aid has increasingly reflected strategic considerations. For example, since the terrorist attacks of September 2001, the objective of reducing terrorism has been of increasing interest among donors giving aid to developing nations. The authors explain the rationale for providing such aid. (JEL F13)

Federal Reserve Bank of St. Louis *Review*, July/August 2013, *95*(4), pp. 327-36.

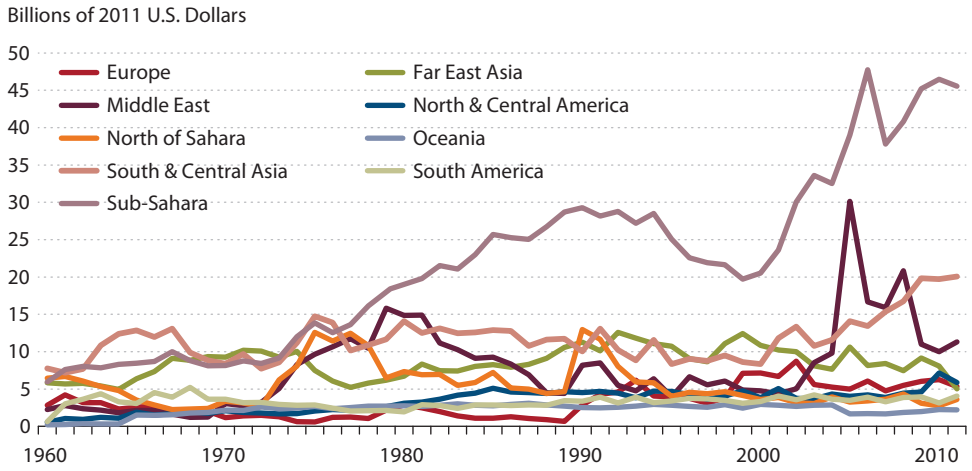
The motives of donor nations for providing foreign aid have evolved over time.[1] For example, prior to World War I, U.S. government-sponsored foreign aid was almost nonexistent. The United States did not financially intervene until World War II, when it began to provide aid abroad, largely through wartime aid and reconstruction efforts. This aid, which focused on "creating markets for the United States by reducing poverty and increasing production in developing countries," also concentrated on "diminishing the threat of communism by helping countries prosper under capitalism" (United States Agency for International Development [USAID, (a)]). These dual goals—humanitarian and strategic—continue to exist in the agency's foreign aid goals to "[develop] the markets of the future…[and] partnerships with countries committed to enabling the private sector investment that is the basis of sustained economic growth to open new markets for American goods" (USAID, (b)). Other nations have similarly altered their donor behavior according to their own histories and evolving environments. For example, a large portion of French aid has been directed to its former colonies, while geopolitical interests have led Japan to give aid to nations broadly supportive of Japanese interests in international forums.

Geopolitical dynamics have evolved over time and, in turn, have affected the decisionmaking processes of donor nations. Figure 1 shows the flow of foreign aid in constant 2011 dollars to

Subhayu Bandyopadhyay is a research officer and economist and E. Katarina Vermann is a senior research associate at the Federal Reserve Bank of St. Louis. The authors thank Alejandro Badel, Maria Canon, and Bill Gavin for their helpful comments.

Figure 1

Net Official Development Assistance to Regions (1960-2011)



NOTE: The figure shows the real value of net official development assistance allocated to specific geographical regions. Countries in sub-Saharan Africa have consistently received the most official development assistance from all donors, and this amount has grown over time. For the majority of the other regions, however, the amounts received have remained fairly constant. The only exceptions are the Middle East and South and Central Asia, where levels of official development assistance have increased dramatically since 2004.

SOURCE: OECD.

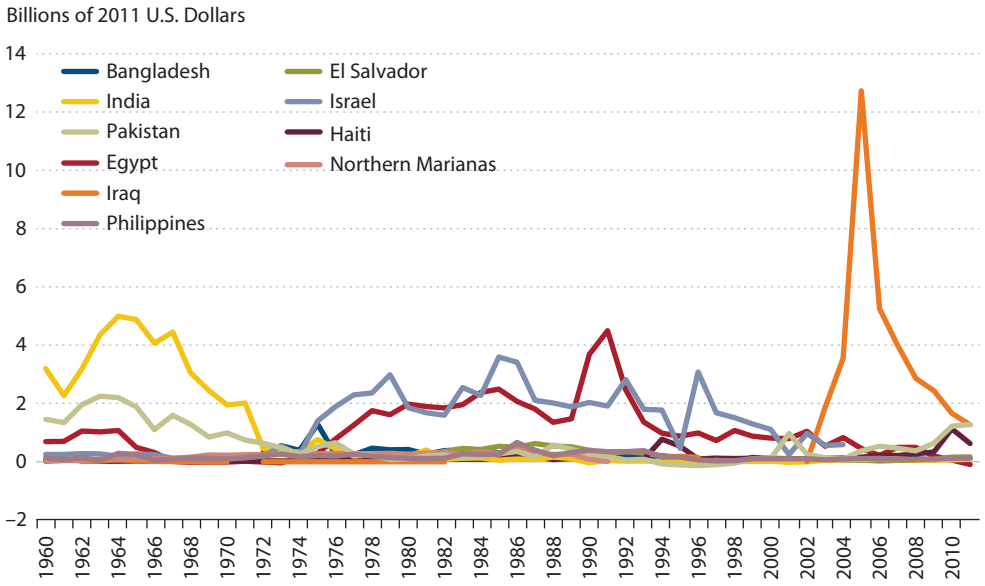
different regions of the developing world from 1960 to 2011. It is clear that the aid flows to sub-Saharan nations have surged upward from approximately $6 billion in 1960 to $46 billion in 2011. This pattern is consistent with the objective of providing foreign aid to alleviate poverty and foster development in the poorest of nations, many of which are located in the sub-Saharan region. Also, there was a spike in foreign aid to Middle Eastern nations in 2005, which has been associated with the Iraq War. The growing involvement of the member nations of the Organisation for Economic Co-operation and Development (OECD) in addressing security concerns originating from the Indian subcontinent and surrounding areas is apparent in the pronounced upward trend in aid flows to the South and Central Asia region.

Figure 2 looks specifically at U.S. aid to specific developing nations. In the 1960s, India, then one of the poorest nations in terms of per capita gross domestic product (GDP), received a disproportionate amount of U.S. aid. This development focus shifts somewhat in the 1970s, when Israel and Egypt—U.S. partners in the Middle East peace process—began receiving more U.S. aid. During the Iraq War in the 2000s, the United States provided large amounts of aid to Iraq.

The objective of this article is to provide a survey of the literature on donor motives and the quantitative impacts of such motives. The next section discusses some theoretical models on the topic. Some of the literature cited in this section also tests the empirical implications of the models. The following section provides a summary of the empirical literature. The final section provides our conclusions.

Figure 2

Total U.S. Official Development Assistance to Top Five Countries in Each Period



NOTES: The figure shows the official development assistance giving patterns to recipient nations that were either in the top five overall recipients and/or in the top five recipients during the Cold War era (1977-90), the interwar period (1991-2001), or the War on Terror (2001-11). The figure shows that, over time, the United States has had a clear preference toward providing aid to Israel and Egypt. Earlier in its history of giving official development assistance, the United States provided substantial aid to India, but that level tapered off while the United States provided substantial aid to Iraq during the War on Terror.

SOURCE: OECD.

DONOR MOTIVES: THEORETICAL MODELS AND SUPPORTING EVIDENCE

In one of the seminal papers on motives behind foreign aid, Dudley and Montmarquette (1976) identify three donor motives. First, donors expect recipient nations to express gratitude in the form of support for donors' interests, perhaps in the sphere of international politics. Second, recipients may trade more with donors, furthering the donors' economic interests. Finally, donors may care that their aid helps residents of recipient nations enjoy a better standard of living. Dudley and Montmarquette (1976) capture these ideas through the following objective function U of the decisionmaker in the donor nations:

$$U = f(X, H), \; H = \sum_{j=1}^{m} H_j, \tag{1}$$

where X is the donor nation's consumption of a private good and H_j is the consumption of the subjectively measured impact of foreign aid to nation j ($j = 1,2,\ldots m$). Further, they assume the following functional form for H_j:

$$H_j = \left(n_j\right)^{\alpha}\left(\frac{a_j}{y_j}\right)^{\gamma},\ 0 \le \alpha \le 1,\ \text{and}\ 0 \le \gamma < 1, \tag{2}$$

where n_j is the population in the recipient nation j, a_j is the per capita aid received by the recipient nation j, and y_j is the per capita gross national product (GNP) of the recipient nation j. Thus, the impact perceived by the donor nation of its aid to another nation is assumed to be increasing in the following characteristics of the recipient nation: population, the per capita amount of aid received, and the per capita GNP. The donor's budget constraint is

$$X + \sum_{j=1}^{m} n_j a_j = Y, \tag{3}$$

where Y is the donor nation's GNP used to (i) consume the private good X and (ii) give foreign aid to the m recipient nations. Using equations (1) through (3), the donor nation's constrained utility maximization problem yields the per capita aid to nation j as

$$a_j = \left[\frac{\gamma k}{y_j^{\gamma} n_j^{1-\alpha}}\right]^{1/(1-\gamma)},\ j\ \ 1,2,...\mathrm{m}, \tag{4}$$

where k is the marginal rate of substitution between H and X (reflecting the value to the donor of an additional unit of H in terms of units of X). Equation (4) suggests that per capita aid to a nation j rises when (i) the value to the donor of giving foreign aid (k) is higher, (ii) recipient nation j's per capita income (y_j) is lower, and (iii) the population of nation j (n_j) is lower.

Dudley and Montmarquette (1976) extend the basic model to allow for administrative costs of aid, which rise less than proportionally to the amount of aid disbursed to a particular recipient nation. This leads to scale economies in giving per capita aid to a nation and thereby tempers the finding of equation (4), which suggests that smaller economies should receive more per capita aid. Using OECD aid commitment data from 1970, Dudley and Montmarquette find broad empirical support for their findings that lower per capita income of the aid recipients tends to raise per capita aid by the donors, while the population of the recipient nation has a more ambiguous effect.

Dudley (1979) considers the interactions between different donors in aid giving. He classifies interactions into two potential categories. First, if the foreign aid from other donor nations adds to the perceived impact of aid giving for a donor nation j, then he classifies such aid as an *international public good*. Second, if aid from other nations reduces the aforementioned impact, then such aid is classified as a *national public good with interaction*. He shows that the first type of aid is associated with a classic public goods problem of free-riding, where smaller nations tend to free-ride on the aid from the larger donors. The intuition is that since larger nations have a lower per capita cost for providing a certain amount of aid, they will be major providers of aid. Given positive spillovers from such aid, the net marginal benefit is lower for smaller donor nations, leading them to provide less per capita aid than larger nations. In the case of *national public goods with interaction*, the donor nations compete for influence through their aid. The

result is a positively sloped reaction function for a nation with respect to aid provided by a rival nation. In equilibrium, foreign aid by each of the donor nations exceeds the level that would be obtained in the absence of such strategic competition. Empirical results of Dudley's (1979) article support the *national public goods with interaction* model, suggesting that the relevant group of OECD nations was engaged in some form of competition for influence.

Lahiri and Raimondos-Møller (2000) provide a political economy framework in which different ethnic groups in a donor nation provide political contributions to the donor government to lobby for aid to the respective source nations of these ethnic groups. While greater lobbying by an ethnic group leads to a greater proportion of aid to the relevant source nation, the corruptibility of the donor government, among other factors, complements such aid flows. In particular, Lahiri and Raimondos-Møller (2000) define corruption as higher when the government tends to enjoy relatively greater marginal utility from political contributions vis-à-vis national welfare. Using this definition, they show that when a richer ethnic group lobbies for a richer recipient nation, an increase in corruption decreases the proportion of aid to the poorer recipient nation. This line of analysis adds an important caveat to the studies discussed above: While the poverty of recipient nations may be an important motive in aid allocation, political economy factors such as ethnic lobbying may also influence the actual aid allocation decision.

Chong and Gradstein (2008) offer a median voter model of foreign aid, where individuals choose to contribute privately and also pay a tax on their income, which finances aid provided by their government. The aggregate aid given is the sum of private donations and official aid, where the latter is financed by income tax revenues raised for that purpose. Thus, aggregate aid is a public good for the citizens of the donor nation. Chong and Gradstein (2008) determine that the aid given is reduced by greater income inequality in the donor nation. They also show that the amount of aid given decreases with the inefficiency of the donor government in effectively providing aid to developing nations. This decrease arises from (i) a reduced willingness on the part of voters to finance inappropriate donor aid and (ii) the direct effects of inefficiencies in aid giving. The authors' empirical results are broadly consistent with their theoretical findings.

Gaytan-Fregoso and Lahiri (2000) present a model that analyzes the effect of foreign aid on illegal immigration from the perspective of a family with a migrant member. Such an analysis is useful for aid decisions because it provides information on how aid may be used for various policy goals such as containing illegal immigration. In this model, aid has two effects. First, aid narrows the gap between incomes in the host and source nations of immigrants, dissuading migration. Second, aid reduces the recipient family's marginal utility of income and hence its perceived utility cost of migration. This second effect tends to encourage migration, which leads to some ambiguity on the effects of aid on migration flows. The authors show that if the volume of initial aid is large enough, then further increases in aid must reduce immigration.

Bandyopadhyay, Sandler, and Younas (2011) evaluate the effect of aid on terrorism. They show that it may be in the interest of developed nations to provide foreign aid to developing nations to help alleviate terrorism-related threats. In their model, conditional aid is tied to the counterterrorism efforts of the aid-recipient nation. Such aid can help neutralize the capabilities of terrorist organizations at their source, leading to a global reduction in terrorism. However, such conditional aid can also lead to popular discontent in recipient nations, especially if their

people feel that their country is being paid to fight the donor nation's war. In such a situation, unconditional aid may help by improving the living conditions of the population, thereby improving the recipient government's perception among its people. The analysis identifies the factors that determine optimal allocation of aid between such conditional and unconditional uses.

Azam and Thelen (2012) present an analysis in which the number of terrorism incidents in a particular nation differs from the number produced because terrorist attacks may be imported. Thus, if enforcement funded by aid contains domestic terrorist groups, the vacuum may be filled by terrorism originating abroad, with a muted response of terrorist attacks to counterterrorism efforts of the country in question. Although the contexts of the aforementioned studies differ, they complement our understanding of how foreign aid may affect counterterrorism efforts and, in turn, the extent of terrorist threats in a particular donor or recipient nation.

DONOR MOTIVES: THE EMPIRICAL LITERATURE

Trumbull and Wall (1994) extend the Dudley and Montmarquette (1976) model by allowing donors to assign different weights to aid impact for different recipient nations. These weights are assumed to differ among recipients to capture, among other factors, the historical, strategic, and geographical differences among recipient nations as perceived by the donors. Allowing for such recipient effects in addition to period effects (i) reduces the role of the recipient's per capita income in determining per capita aid flows and (ii) amplifies the role of infant mortality (a measure of need) and civil rights in recipient nations. Simply put, a poorer nation may not necessarily receive more aid from an altruistic donor because the donor may punish the recipient government for political oppression. Along similar lines, Bandyopadhyay and Wall (2007) find that for an average recipient nation, a one-standard-deviation increase in civil/political rights and government effectiveness corresponds to increases of $29 million and $54 million, respectively, in total aid received by these nations.

Alesina and Dollar (2000) study donor behavior for individual nations using OECD data on bilateral aid. While they find broad support for the importance of the per capita income of recipient nations in receiving aid, they also find that ties to former colonies or strategic factors are quite significant. For example, a huge fraction of U.S. aid goes to Israel and Egypt, which are not among the poorer recipient nations. France's aid is concentrated on its former colonies, while more of Japan's aid goes to nations with comparable international political interests, as measured by similarity in U.N. voting patterns. The study also finds that donor countries reward developing countries pursuing greater democratization. Aid allocation surges to nations that liberalize in this dimension, although economic liberalization is not significantly associated with greater aid flows.

Like Alesina and Dollar (2000), more recent studies find that international political factors are important in aid giving. Kuziemko and Werker (2006) and Dreher, Strum, and Vreeland (2009) find a large positive effect of U.N. Security Council membership on aid recipients. These amounts increase in years when (i) the Security Council and/or the United Nations receive higher levels of media coverage (indicative of years with tumultuous foreign affairs) and (ii) major international events occur. Further, the levels of aid allocated to member countries of the U.N. Security

Council "sharply increase in the year in which a country is elected to the Security Council, remain high throughout the two-year term, and return to their earlier level almost immediately upon completion of the term" (Kuziemko and Werker, 2006, p. 907). Hoeffler and Outram (2011) find that aid also increases with an increase in the correlation between the U.N. votes of a donor and recipient country, though this correlation is sensitive to specification.

Domestic politics is also quite important for aid giving. Milner and Tingley (2010) examine five types of foreign aid policy votes in the U.S. House of Representatives from 1979 to 2003. They find that votes on aid correlate with the material interests of representative constituents. Specifically, they find that voting districts with relatively higher capital endowments are more supportive of foreign aid, following the Stolper-Samuelson theorem.[2] Also, they find broad political support for food aid and military aid where they find no stark distributional effects between labor and capital. Political ideology also seems to play a role. For example, using data for a panel of donor countries from 1971 to 2002, Tingley (2010) finds that the more economically conservative the ideological orientation of the donor government, the less money there is in the overall aid budget.

Donors may use foreign aid to foster trade with the recipient country. If aid to a recipient nation is used by a developing nation to buy donor nation exports, it can act as an export promotion strategy for the developed nation. The potential benefits for the donor include terms-of-trade benefits and larger real incomes for factors used intensively in the export good. Following this line of reasoning, Younas (2008) finds that "a substantially larger amount of aid is provided to recipients who import capital goods, while imports by other category groups have no significant effects. Given that developed donor nations are major producers and exporters of capital goods, this result at least partially supports their trade benefits motive" (p. 661). Other authors also find that trade interest—generally measured as imports from and exports to a recipient country as a percent of donor GDP—is correlated with higher levels of aid (Nath and Sobhee, 2007; Hoeffler and Outram, 2011; Dietrich, 2012), though these estimates may be sensitive to specification (Hoeffler and Outram, 2011).

Finkel, Pérez-Liñán, and Seligson (2007) argue there are two mechanisms through which foreign aid can promote democracy in developing nations. First, foreign aid can indirectly promote democracy "by transforming some of the structural conditions that serve as prerequisites for regime transition or survival" (p. 410). To illustrate, foreign aid "may promote modernization, encourage better economic performance, and foster class transformations, all of which may have long-term implications for democratic development" (p. 410). Second, aid can directly promote democracy "by empowering agents…that struggle for regime change in the domestic arena" (p. 410) with aid targeted toward education, political parties, labor unions, human rights groups, and women's advocacy networks. Their empirical work indicates that an increase in a donor's democracy and/or government-targeted aid increases the aid recipient's level of democracy. Bermeo (2011) confirms this result with logit regressions and Monte Carlo simulations and finds that, holding all control variables constant at their median, "a change in the amount of democratic aid from its value at the 25th percentile to its value at the 75th percentile is associated with an increase in the likelihood of democratic transition in a given country during a given year from a baseline of 2.15 to 3.12 percent, a 45 percent increase" (p. 2025).

Another recipient characteristic that may influence aid is the presence of conflict in the recipient nation. Balla and Reinhardt (2008) argue that "the donor's perception of the conflict [in the recipient country] will determine how useful it expects its aid to be at achieving [the donor's] interests. In light of the propensity of the conflict to alter aid's utility, the existence of the conflict may cause the donor to alter the amount, or the existence, of that aid" (p. 2568). Evidence from this article indicates that, for the 22 donors examined, 14 are more likely to allocate additional aid to countries experiencing conflict, 14 are more likely to allocate aid to a region experiencing conflict, and 9 are more likely to reduce the amount of aid allocated to a recipient as conflict increases. In an assessment of conflict and aid in a panel of 122 recipients from 1960 to 1997, Balla and Reinhardt (2008) find that proximity to conflict is statistically significant and positively correlated with respect to aid allocations. Further, they find that the likelihood for a donor to select an aid recipient increases as conflict increases in the recipient nation.

In recent years, national security has been in the forefront of foreign policy concerns. As our theoretical section outlines, it is possible to use aid to reduce threats from terrorism originating in developing nations. According to Fleck and Kilby (2010, p. 185), "starting in the mid 1990s and continuing into the War on Terror, the importance of need as a criterion for aid eligibility fell so that the probability that a higher income (less poor) developing country would receive aid increased over time."[3] They find that higher-income recipients were 9 percent more likely to receive aid during the War on Terror than during the Cold War era. Demierl-Pegg and Moskowitz (2009) confirm this result: They find that economic development (per capita GDP) had a larger impact on the probability of receiving foreign aid during the Cold War than strategic interests (bilateral trade and political alliance) and that per capita GDP also outweighed bilateral trade and regime type in aid allocation. Further, the impact of economic development and human rights on the probability of a country receiving aid and the amount of aid to that country were higher during the 1990s than during the Cold War era. Indeed, aid budgets overall increased among 22 donor countries since the start of the War on Terror (Dreher and Fuchs, 2011), and aid allocation decisions changed during this period after controlling for natural disasters, recipient country size, recipient country movement to democracy, and U.N. votes. Young and Findley (2011) find that targeted (conditional) aid can be an effective tool to deter terrorism, especially when directed at education, health, civil society, and conflict prevention in the recipient country.

CONCLUSION

As many poorer nations started on their path of development in the postcolonial period, former colonial powers tried to facilitate this process through provision of resources in the form of foreign aid. This motive was central in foreign aid flows throughout the 1960s and still remains one of the multiple objectives that determine aid. In addition to pure altruism, strategic interests have always been relevant. In the Cold War era, aid was used as an instrument to buy allegiance. More recently, aid has focused substantially on security concerns, marked especially by the watershed event of September 2001. This review has attempted to provide a summary of some important theoretical models and empirical findings proposed in the literature to further critical thinking about these issues.

NOTES

[1] We take our definition of foreign aid from Perkins, Radelet, and Lindauer (2006, p. 521): "Foreign aid consists of financial flows, technical assistance, and commodities given by the residents of one country to the residents of another country, either as grants or as subsidized loans." Precise characterization of what is and what is not foreign aid is more nuanced.

[2] If aid fosters domestic production of a certain good (either directly through production of some U.S. good that may be given as aid or indirectly through demand for a U.S. good from aid-receiving nations), then Stolper-Samuleson effects will predict that the real income of factors intensive in these goods will rise. Given that the United States is relatively capital abundant compared with most of the poorer, labor-abundant developing nations, Stolper-Samuelson effects would suggest that U.S. exports will raise the real income of capital.

[3] This analysis excluded countries such as Egypt, Israel, Iraq, and Afghanistan from analysis. These countries exhibit high levels of strategic importance but were excluded because of either a lack of data (Iraq and Afghanistan) or because they were extreme outliers (Egypt and Israel).

REFERENCES

Alesina, Alberto and Dollar, David. "Who Gives Foreign Aid to Whom and Why?" *Journal of Economic Growth*, March 2000, *5*(1), pp. 33-63.

Azam, Jean-Paul and Thelen, Véronique. "Where to Spend Foreign Aid to Counter Terrorism." Working Paper No. 725, Institut d'Économie Industrielle, June 2012; http://idei.fr/doc/wp/2012/wp_idei_725.pdf.

Balla, Eliana and Reinhardt, Gina Yannitell. "Giving and Receiving Foreign Aid: Does Conflict Count?" *World Development*, December 2008, *36*(12), pp. 2566-85.

Bandyopadhyay, Subhayu; Sandler, Todd and Younas, Javed. "Foreign Aid as Counterterrorism Policy." *Oxford Economic Papers*, July 2011, *63*(3), pp. 423-47.

Bandyopadhyay, Subhayu and Wall, Howard J. "The Determinants of Aid in the Post-Cold War Era." Federal Reserve Bank of St. Louis *Review*, November/December 2007, *89*(6), pp. 533-47; http://research.stlouisfed.org/publications/review/07/11/Bandyopadhyay.pdf.

Bermeo, Sarah B. "Foreign Aid and Regime Change: A Role for Donor Intent." *World Development*, November 2011, *39*(11), pp. 2012-31.

Chong, Alberto and Gradstein, Mark. "What Determines Foreign Aid? The Donors' Perspectives." *Journal of Development Economics*, August 2008, *87*(1), pp. 1-13.

Demirel-Pegg, Tijen and Moskowitz, James. "U.S. Aid Allocation: The Nexus of Human Rights, Democracy, and Development." *Journal of Peace Research*, March 2009, *46*(2), pp. 181-98.

Dietrich, Simone. "Bypass or Engage? Explaining Donor Delivery Tactics in Foreign Aid Allocation." *International Studies Quarterly*, 2012; early view, doi: 10.1111/isqu.12041.

Dreher, Axel and Fuchs, Andreas. "Does Terror Increase Aid?" *Public Choice*, December 2011, *149*(3-4), pp. 337-63.

Dreher, Axel; Sturm, Jan-Egbert and Vreeland, James Raymond. "Development Aid and International Politics: Does Membership on the UN Security Council Influence World Bank Decisions?" *Journal of Development Economics*, January 2009, *88*(1), pp. 1-18.

Dudley, Leonard M. "Foreign Aid and the Theory of Alliances." *Review of Economics and Statistics*, November 1979, *61*(4), pp. 564-71.

Dudley, Leonard M. and Montmarquette, Claude. "A Model of the Supply of Bilateral Foreign Aid." *American Economic Review*, March 1976, *66*(1), pp. 132-42.

Finkel, Steven E.; Pérez-Liñán, Aníbal S. and Seligson, Mitchell A. "The Effects of U.S. Foreign Assistance on Democracy Building, 1990-2003." *World Politics*, April 2007, *59*(3), pp. 404-39.

Fleck, Robert K. and Kilby, Christopher. "Changing Aid Regimes? U.S. Foreign Aid from the Cold War to the War on Terror." *Journal of Development Economics*, March 2010, *91*(2), pp. 185-97.

Gaytan-Fregoso, Helena and Lahiri, Sajal. "Foreign Aid and Illegal Immigration." *Journal of Development Economics*, December 2000, *63*(2), pp. 515-27.

Hoeffler, Anke and Outram, Verity. "Need, Merit or Self-Interest—What Determines the Allocation of Aid?" *Review of Development Economics*, May 2011, *15*(2), pp. 237-50.

Kuziemko, Ilyana and Werker, Eric D. "How Much Is a Seat on the Security Council Worth? Foreign Aid and Bribery at the United Nations." *Journal of Political Economy*, October 2006, *114*(5), pp. 905-30.

Lahiri, Sajal and Raimondos-Møller, Pascalis. "Lobbying by Ethnic Groups and Aid Allocation." *Economic Journal*, March 2000, *110*(462), pp. C62-79.

Milner, Helen V. and Tingley, Dustin H. "The Political Economy of U.S. Foreign Aid: American Legislators and the Domestic Politics of Aid." *Economics and Politics*, July 2010, *22*(2), pp. 200-32.

Nath, Shyam and Sobhee, Sanjeev K. "Aid Motivation and Donor Behavior." *American Review of Political Economy*, 2007, *5*(1), pp. 1-13.

Perkins, Dwight H.; Radelet, Steven and Lindauer, David L. *Economics of Development*. Sixth Edition. New York: W.W. Norton, 2006.

Tingley, Dustin H. "Donors and Domestic Politics: Political Influences on Foreign Aid Effort." *Quarterly Review of Economics and Finance*, February 2010, *50*(1), pp. 40-49.

Trumbull, William N. and Wall, Howard J. "Estimating Aid-Allocation Criteria with Panel Data." *Economic Journal*, July 1994, *104*(425), pp. 876-82.

United States Agency for International Development (a). "USAID History." Undated; http://www.usaid.gov/who-we-are/usaid-history.

United States Agency for International Development (b). "What We Do." Undated; http://www.usaid.gov/what-we-do.

Younas, Javed. "Motivation for Bilateral Aid Allocation: Altruism or Trade Benefits." *European Journal of Political Economy*, September 2008, *24*(3), pp. 661-74.

Young, Joseph K. and Findley, Michael G. "Can Peace Be Purchased? A Sectoral-Level Analysis of Aid's Influence on Transnational Terrorism." *Public Choice*, December 2011, *149*(3-4), pp. 365-81.