Jeffrey Wurgler
New York University

Ekaterina Zhuravskaya
Centre for Economic and Financial Research (Moscow) and Centre for Economic Policy Research

Does Arbitrage Flatten Demand Curves for Stocks?*

In textbook theory, demand curves for stocks are kept flat by riskless arbitrage between perfect substitutes. In reality, however, individual stocks do not have perfect substitutes. We develop a simple model of demand curves for stocks in which the risk inherent in arbitrage between imperfect substitutes deters risk-averse arbitrageurs from flattening demand curves. Consistent with the model, stocks without close substitutes experience higher price jumps upon inclusion into the S&P 500 Index. The results suggest that arbitrage is weaker and mispricing is likely to be more frequent and more severe among stocks without close substitutes.

I. Introduction

In textbook theory, demand curves for stocks are kept flat by arbitrage between perfect substitutes. Myron Scholes argues in his study of large-block sales that "the market will price assets such that the expected rates of return on assets of similar risk are equal. If any particular asset should be selling to yield a higher expected return due solely to the increase in the quantity of shares outstanding . . . investors seeing these profit opportunities would soon arbitrage them away" (1972, p. 182). Arbitrage is possible because "the shares a firm sells are not unique works of art but abstract rights to an uncertain income stream for which close counterparts exist either directly or indirectly via combinations of assets of various kinds" (p. 179).

The arbitrage Scholes envisions is fundamental in finance theory: buy the underpriced stock and simultaneously short a perfect substitute or do the opposite

* We thank Marianne Bertrand, Ravi Dhar, Ken French, Will Goetzmann, Sergei Guriev, Brian Hall, Vassil Konstantinov, Andrew Metrick, Randall Morck, Sendhil Mullainathan, Mike Rashes, Geert Rouwenhorst, Ivo Welch, participants of seminars at Harvard University, Yale University, Lehman Brothers, the National Bureau of Economic Research, the 2000 Western Finance Association meeting, an anonymous referee, and especially Malcolm Baker, John Campbell, Kristian Rydqvist, and Andrei Shleifer for helpful comments. We also thank James Choi, Rick Mendenhall, Andrei Shleifer, Matt Van Vlack, and Standard and Poor's for providing data.

(*Journal of Business,* 2002, vol. 75, no. 4)

0021-9398/2002/7504-0002$10.00

when faced with an overpriced stock. The absence of such arbitrage opportunities leads to the arbitrage pricing theory, the capital asset pricing model, net present value rules, and the Modigliani-Miller theorems. The argument is appealing because the Scholes trade requires no net investment and earns an instant and riskless profit from a portfolio that can be liquidated whenever prices move back in line. The argument is powerful because competition between Scholes arbitrageurs keeps demand curves for individual stocks flat and therefore insulates stock prices from shifts in demand and supply unrelated to fundamental value. The Scholes arbitrage is one of the most fundamental mechanisms in the theory of efficient markets.

Is it really reasonable to think that arbitrage works so well? To lend some empirical substance to the argument, one can point to the fact that between-stock arbitrage is widely pursued by investment professionals. Barr Rosenberg, D. E. Shaw, and Long-Term Capital (Spiro 1998) are well-known examples. *The U.S. Offshore Funds Directory* (1999) lists many dozens of offshore hedge funds that use "pairs trading" or "relative value trading" or "market neutral long/short trading" as one of their principal equity investment strategies.[1]

Although there is undoubtedly some arbitrage activity in real stock markets, it seems unlikely to work as effectively as the textbook theory suggests. One big problem is that individual stocks do not, in practice, have perfect substitutes. Would-be arbitrageurs who take positions in a mispriced stock and hedge with opposite positions in imperfect substitutes must bear the "arbitrage risk" that the two return streams do not cancel out.[2] Risk averse arbitrageurs will trade less aggressively if they must bear this arbitrage risk.

These contrasting observations motivate the need for a closer examination of the argument for flat demand curves for stocks. In this article, we start with a simple model of demand curves for stocks that captures the theoretical role of arbitrage and shows its limits. We then test some predictions of this model using the cross section of price responses of stocks added to the Standard and Poor's 500 Index.

The model considers two groups of investors, arbitrageurs and nonarbitrageurs. Arbitrageurs differ from nonarbitrageurs in two ways: they hold zero-net-investment portfolios ("arbitrage portfolios" that require no capital up front), and they have correct beliefs about the long-run fundamental value of stocks. Arbitrageurs are assumed to be undiversified and thus averse to the arbitrage risk associated with pairs trades between imperfect substitutes. We horizontally sum the arbitrageurs' and nonarbitrageurs' demand curves to get an aggregate demand curve. This curve's slope depends on four factors in intuitive directions. It is flatter when (i) the stock has closer substitutes (less arbitrage risk), (ii) the risk aversion of arbitrageurs is lower, (iii) the heter-

1. Gatev, Goetzmann, and Rouwenhorst (2000) examine the investment appeal of relative value trading strategies. Richards (2000) studies the risk of these strategies.

2. Jensen (1994, p. 50) calls this notion of arbitrage "risk arbitrage": "Arbitrage in the financial markets, more accurately known as risk arbitrage because it does not usually involve the simultaneous purchase and sale of the *same* item" (emphasis by the author).

ogeneity of nonarbitrageurs' beliefs is lower, and (iv) the number of arbitrageurs is higher. In various limiting cases, arbitrage completely flattens the demand curve.

This model makes the straightforward prediction that the price response to an excess demand shock increases with the shock's size and the stock's arbitrage risk. We test this using the cross section of price responses of stocks added to the S&P 500 Index. Upon addition to the index, stocks are bought heavily by index funds that track the index. Since a stock is added to the index only when another stock must be dropped (usually because of merger, takeover, or bankruptcy), the sudden surge in index fund demand for the added stock is arguably an exogenous demand shock. We focus on additions made between 1976 and September 1989, since, more recently, index changes are preannounced. This causes a serious confounding effect since arbitrageurs will very likely prefer front-running the index funds (between the announcement date and the effective date) over the Scholes arbitrage trade (Beneish and Whaley 1996; Ip 1998).

Harris and Gurel (1986) and Shleifer (1986) report that added stocks' prices jump an average of 3% on the day of the change. Shleifer argues that the result reflects downward-sloping demand curves. Other explanations have been suggested. Perhaps S&P 500 addition really does reflect good news about the prospects of the company, despite S&P's claims to the contrary.[3] But this does not explain why the addition effect has grown strongly over time, parallel to the growth in index funds. Another possible explanation of the addition effect is that the price response capitalizes a reduction in future trading costs because of the liquidity associated with institutional buyers. However, Beneish and Whaley (1996) find that bid-ask spreads are not permanently reduced following S&P 500 addition. Kaul, Mehrotra, and Morck (2000) also find no permanent reduction in bid-ask spreads in their Toronto Stock Exchange 300 index-reweighting experiment. In addition, since index funds buy and hold, they reduce the public float. This reduces, not increases, liquidity.

For each stock added to the S&P 500, we measure the size of the index fund demand shock and the stock's arbitrage risk. Index fund size data is from Standard and Poor's. Arbitrage risk is measured as the historical variance of a zero-net-investment portfolio that holds \$1 long (equivalently, short) in the added stock and \$1 short (equivalently, long) in a portfolio of substitutes. This intuitive measure of arbitrage risk falls out of the model.

The main empirical results are as follows. First, individual stocks do not have perfect substitutes. While this is obvious in general, it is surprising how hard it is to find even reasonably close substitutes. For our median stock, we are not able to find substitutes that could hedge away even a quarter of the

3. Standard and Poor's press releases announcing index changes carry the boilerplate note: "Company additions to and deletions from S&P equity indexes do not in any way reflect an opinion on the investment merits of the company."

stock's daily return variance.[4] Second, as predicted by the model, high-arbitrage risk stocks do indeed experience higher price jumps, controlling for demand shock size. Third, again as predicted, stocks hit by especially large index fund demand shocks also experience higher price jumps, controlling for arbitrage risk. Finally, our regressions provide some numerical estimates of the slopes of stocks' demand curves. Our elasticity estimates from the S&P 500 addition context are in the range of previous estimates. For reference, we provide a summary of our own and previous elasticity estimates in a table.

Our results are consistent with the hypothesis that risk prevents arbitrage from completely flattening demand curves for stocks. The results suggest that typical between-stock arbitrage trades are too risky to be worth a couple of percentage points, which is what arbitrageurs stand to gain from reversion of the average S&P effect in our sample. However, the arbitrage argument does explain some of the cross section: stocks that have closer substitutes have flatter demand curves.

The long-run effect of S&P 500 additions also bears on our model. Unfortunately, the exercise of measuring long-run returns is hampered by the lack of agreement on how to adjust for risk, a small sample, and the variability of cumulative abnormal returns. Shleifer (1986), Dhillon and Johnson (1991), Beneish and Whaley (1996), and Lynch and Mendenhall (1997) find that the price jumps partially revert (to varying degrees) but have a substantial permanent component (where permanent is taken to mean not reversing, on average, within a few months), while Harris and Gurel (1986) find that the effect completely reverts within a short 2 weeks. We have been able to verify prior results but, given no generally accepted way to adjust for risk, we do not investigate the long-run effect issue anew. All prior authors agree that at least a portion of the effect reverts in the long run. In our model, this is the return that attracts arbitrageurs in the first place.

Arbitrage risk may shed light on a number of studies that suggest demand curves for stocks slope downward. Kaul et al. (2000) document price changes following a reweighting of the Toronto Stock Exchange 300 Index. Asquith and Mullins (1986) document negative announcement returns for primary equity offerings, and more negative returns for larger offerings. Analogous results are found by Bradley, Desai, and Kim (1988) for tender offers; Holthausen, Leftwich, and Mayers (1990) for buyer-initiated block trades; and Loderer, Cooney, and Van Drunen (1991) for primary offerings by regulated firms. To the extent that these results cannot be explained by new information or changes in liquidity, they are inconsistent with the textbook assumption of flat demand curves. Between-stock arbitrage trades would be possible in all

4. This agrees with Roll (1988). He finds that, across stocks, the average R^2 for daily returns is about .20, controlling for other stocks' returns, fundamental factors, or industry or market returns.

of these settings, and our results suggest that a closer study of the risks of such arbitrage may help to explain the cross section of these results.

More generally, the results suggest that arbitrage risk can be added to a growing list of forces that inhibit market efficiency. This list includes informational and transaction costs (Stoll 2000), heterogeneous beliefs about fundamental value (Bagwell 1992), noise trader sentiment risk (De Long et al. 1990), short-sales constraints (Chen, Hong, and Stein 2002), and agency costs in delegated fund management (Shleifer and Vishny 1997).

As any of these constraints on arbitrage bind more strongly, mispricings become more likely and more severe. For example, our results suggest that high-arbitrage risk stocks have steeper demand curves. This means that such stocks would be particularly likely to exhibit anomalies driven by uninformed changes in supply and demand. Some initial support for this broader conjecture comes from our finding that small stocks are, in general, riskier to arbitrage than large stocks. As Loughran and Ritter (2000, p. 363) observe, "just about every known stock market pattern is stronger for small firms than for large firms."

This article proceeds as follows. In Section II, we present a simple model in which risk deters arbitrageurs from flattening aggregate demand curves. In Section III, we look at the arbitrage risk of stocks added to the S&P 500 and regression results that suggest that arbitrage risk affects the slope of demand curves. Section IV concludes.

II. A Model of Demand Curves for Stocks

The model consists of two types of investors, arbitrageurs and nonarbitrageurs. They differ in two dimensions. First, arbitrageurs have correct and homogeneous beliefs about the fundamental value of all assets, while nonarbitrageurs have heterogeneous beliefs. Second, arbitrageurs are subject to a zero-net-investment constraint. This is a defining property of real Wall Street arbitrage strategies such as pairs trading or relative value trading. In this section we model their respective demand curves and horizontally sum them to derive the aggregate demand curve. Then we invert the aggregate demand curve to show the determinants of an asset's price response to a demand shock.

A. Demand Curves of Arbitrageurs

Consider a market of n risky assets, purchased at $t = 0$ and sold at $t = 1$. Returns $\tilde{R}_i = (\tilde{P}_{i1}/P_{i0}) - 1$ are distributed normally. For some reason, left exogenous for now, asset j is mispriced—it has an expected return of $\bar{R}_j \neq 0$, while other assets have expected returns of zero. The risk-averse arbitrageur is lured into the market by this mispricing.

Because textbook arbitrage requires no initial outlay, we model the arbitrageur subject to an initial wealth constraint of $0. The arbitrageur has exponential utility with constant absolute risk aversion k. It is not difficult to show that these assumptions imply a mean-variance objective and the following dollar excess demand for asset j:

$$x_j^a = \frac{\bar{R}_j}{kA_j}, \tag{1}$$

where

$$A_j \equiv \min_{x_i:\ i \neq j} \mathrm{Var}\,(\bar{R}_j + \sum_{i \neq j} \bar{R}_i x_i), \tag{2}$$

subject to

$$\sum_{i \neq j} x_i = -1.$$

This solution is intuitive. The arbitrageur faces a mean-variance objective. Holdings of asset j alone determine the expected return of the portfolio, since only it has nonzero expected return. Given the commitment $\$x_j^a$, all other assets are held to minimize the overall risk of the portfolio, subject only to the zero-net-investment constraint. Quantity A_j is asset j's "arbitrage risk." It is the variance of the minimum variance portfolio, which is long \$1 in asset j and short, on net, \$1 in other assets.[5] For example, if asset j has a perfect substitute, A_j is zero. If the expected return on asset j is positive (negative), the arbitrageur invests $\$x_j^a$ long (short) in asset j and $\$x_j^a$ short (long) across other assets to minimize the total risk of the portfolio. The arbitrageur takes a small position when the potential gains are small, his risk-aversion is high, and substitutes are hard to find.

It is useful to visualize the solution as a downward-sloping demand curve. Denote $P_j^* = E[P_{j1}]$. Under the approximation that $\bar{R}_j \approx 1 - (P_j/P_j^*)$, equation (1) indicates that the arbitrageur's excess demand curve has an intercept of 1 and a slope of $-kA_j$ in $(x, P_j/P_j^*)$ space. As is intuitive, if risk aversion is zero or the asset has perfect substitutes, the demand curve is infinitely elastic and perfectly flat.

B. Demand Curves of Nonarbitrageurs

To define the excess demand of nonarbitrageurs, we propose the specification

$$x_j^{na} = \frac{\bar{R}_j}{h_j}, \tag{3}$$

5. Note the distinction between arbitrage risk, idiosyncratic risk, and nondiversifiable risk. Idiosyncratic risk is usually associated with the error term in asset pricing models. Nondiversifiable risk is closer, but arbitrage imposes the constraint that the hedge position be equal and opposite the position in the mispriced asset (zero-net investment). Therefore, arbitrage risk is weakly greater than nondiversifiable risk.

where h_j is simply a parameter that indexes the heterogeneity of nonarbitrageurs' beliefs about asset j's fundamental value P_j^*. With homogeneous beliefs (small h_j), the excess demand curve of nonarbitrageurs is flat. But if there is wide heterogeneity of beliefs, large changes in price (and large expected returns) are required to induce marginal nonarbitrageurs to modify their demands. Heterogeneity of beliefs about value is convincingly demonstrated by Bagwell (1992) and Kandel, Sarig, and Wohl (1999).

Assuming that nonarbitrageurs' excess demand is zero at the fundamental price, then, using the return approximation as above, their excess demand curve has an intercept of 1 and a slope of $-h_j$ in $(x,\ P_j/P_j^*)$ space.

C. The Aggregate Demand Curve and the Mispricing Induced by a Demand Shock

The aggregate excess demand curve is the horizontal sum $x_j^{agg} = x_j^a + x_j^{na}$. Using the previous results, one can easily show that the aggregate curve is approximately

$$\frac{P_j}{P_j^*} = 1 - \frac{x_j^{agg}}{\frac{1}{kA_j} + \frac{1}{h_j}}. \tag{4}$$

If there are N identical arbitrageurs instead of just one, symmetry implies that N/kA_j replaces $1/kA_j$ in equation (4). Under the assumptions made above, the aggregate demand curve has an intercept of 1 and a slope of $-1/[(N/kA_j) + (1/h_j)]$ in $(x,\ P_j/P_j^*)$ space.

With this model of the aggregate demand curve, it is simple to see how price responds to excess demand shocks. Suppose that a sudden demand shock of \$$S$ hits asset j at time $t = s$, between $t = 0$ and $t = 1$, and the shock is exogenous in the sense that it contains no fundamental news about asset j. For example, the sudden index fund demand following a stock's inclusion into the S&P 500 Index reflects the charter of index funds, not new information about fundamental value. Both arbitrageurs and nonarbitrageurs help to accommodate this shock: $x_{js}^{agg} = x_{js}^a + x_{js}^{na} = -S$. Assuming that asset j was initially correctly priced at time 0, that is, that $P_{j0} = P_j^* = E[P_{j1}]$, we can plug this demand into (4) to show that the excess return to a demand shock of \$$S$ is simply

$$\frac{P_{js}}{P_{j0}} - 1 \approx \frac{S}{\frac{N}{kA_j} + \frac{1}{h_j}}. \tag{5}$$

Special cases illustrate the comparative statics. When the asset has perfect substitutes (A_j is zero), arbitrageurs completely flatten the aggregate demand curve, so shocks are absorbed without any price change. This also happens when there are infinitely many arbitrageurs or when the arbitrageurs are risk neutral. The presumed existence of infinitely many small and locally risk-neutral arbitrageurs is the traditional fallback argument for market perfection and efficiency even when arbitrage is risky.[6] (It is useful to think of N/k as the aggregate risk-bearing capacity of arbitrageurs. This capacity is infinite when there are infinitely many of them or when any of them are risk neutral.) Otherwise, if the available substitutes are imperfect, there are few and risk-averse arbitrageurs, and nonarbitrageurs disagree about the fundamental value of the asset, the aggregate demand curve slopes significantly, and the price response to an excess demand shock is large. A final interesting case is infinite h_j. In this case, arbitrageurs are functionally similar to the market makers modeled by Grossman and Miller (1988). One can verify that these qualitative features of the price response do not depend on the approximation we used to derive equation (5).

A more general remark is in order at this point. Most financial economists have a strong intuition that idiosyncratic risk such as arbitrage risk is irrelevant because it can be diversified away. There are a few reasons to question the hasty application of this intuition in all settings.

First, as a practical matter, arbitrageurs are often not fully diversified. The dramatic experience of Long-Term Capital in the fall of 1998 is sufficient to prove this point: the fund's failure can ultimately be traced to an unfortunate position on Russian debt. By definition, a well-diversified fund could not fail because of a single position. More generally, real arbitrageurs tend to have specialized knowledge that identifies a small number of good bets at one time, and they spread their limited capital and risk-bearing capacity across a small number of positions (Shleifer and Vishny 1997). The assumption that arbitrageurs are, in general, averse to idiosyncratic risk is clearly an abstraction but one that does not seem irrelevant in real financial markets.[7]

Second, as a theoretical matter, there is simply no opportunity to diversify across S&P 500 additions or similar events that occur infrequently over time. There is only one change to the S&P 500 every few weeks, on average, so arbitrageurs cannot hold positions in these mispricings all at once. We develop this point in the conclusion.

III. An Empirical Test Using S&P 500 Additions

One prediction of the model in Section II is that high-arbitrage-risk stocks have steeper aggregate demand curves than low-arbitrage-risk stocks, other things equal. An ideal test of this prediction would select two stocks which differ in their risk levels but are similar in other respects, subject them to identical informationless excess demand shocks, and compare their price re-

6. Shleifer and Vishny (1997) argue that N is effectively small in reality.

7. Green and Rydqvist (1997) find that idiosyncratic risk plays a role in the pricing of Swedish government lottery bonds.

sponses. The stock with the higher arbitrage risk should have a larger price response if its demand curve is indeed steeper.

This section presents empirical tests based on this intuition. We consider the price responses of stocks added to the S&P 500 Composite Index. Upon inclusion into the index, stocks are bought heavily by numerous index funds that track the index. These funds have an incentive to minimize tracking error by buying as soon as the changes are made effective. Since a stock is added to the index only when another stock must be dropped, usually because of merger, takeover, or bankruptcy, the sudden surge in index fund demand for the added stock arguably represents an exogenous demand shock.

Since the supply curve of the stock is fixed and vertical over the relevant horizon, the size of the price jump following this demand shock is determined by the slope of the stock's aggregate demand curve and the size of its shift. If arbitrage risk is a limiting factor in the slope of demand curves, we should find that this jump is differentially larger for high-arbitrage-risk stocks, controlling for shock size. Note that we do not evaluate a trading strategy. Our model predicts that the equilibrium price response to a demand shock depends on the risk of a particular zero-net-investment trading strategy.

Although equation (5) applies to negative shocks and positive shocks, we do not examine index deletions, since they usually happen during periods of very intensive confounding news—usually merger, takeover, or bankruptcy. Goetzmann and Garry (1986), Harris and Gurel (1986), and Lynch and Mendenhall (1997) study small, carefully screened samples of index deletions.

Our tests require measures of arbitrage risk and shock size for each addition event. To measure arbitrage risk, the model suggests that we use the variance of the portfolio that holds long \$1 of the stock and short \$1 in a portfolio of close substitutes. In the S&P 500 addition setting, it is natural to think of shorting the added stock and going long in substitutes, but of course $\mathrm{Var}(R_i - R_j) = \mathrm{Var}(R_j - R_i)$. We consider two different methods of identifying substitutes and characterize the resulting arbitrage risk measures.

To measure shock size, we rely on Standard and Poor's own estimates of the cumulative size of all funds indexed to the S&P 500. Given these estimates, the fact that the S&P 500 is a value-weighted index, the aggregate capitalization of the S&P 500, and the capitalization of the added stock, we can estimate the dollar value of the index fund demand shock to each added stock. We also measure demand shock size as the percent shock (in terms of the added firm's capitalization).

We first describe our addition event sample. Next we detail the construction of variables and present summary statistics. Then we show regressions of the determinants of event day returns and some figures that show that most of the influence of arbitrage risk and shock size on event returns lasts at least 20 trading days, but beyond that it is difficult to make meaningful inferences. We close with some numerical estimates of the slopes of demand curves for stocks implicit in the regression results and compare them with a range of previously published estimates.

A. S&P 500 Index Addition Sample

We consider S&P 500 Index additions between September 22, 1976, and September 30, 1989. Standard and Poor's did not record announcement dates prior to this starting date. Starting in October 1989, Standard and Poor's began to preannounce index changes several days before they were effective. Thus, in the earlier period there is a single "event day" in which the change was announced and immediately made effective.[8] Our setting represents the cleanest experiment in which to study the Scholes argument that we have found.

There were 278 index changes announced between September 22, 1976, and September 30, 1989. Among the added stocks, we drop 18 due to missing data (event-day returns or historical returns needed to compute arbitrage risk measures) and one due to outlying event day returns.[9] Raw event day returns for the remaining 259 stocks in this subsample average 3.29% (SD = 2.32%) and returns over the market average 3.16% (SD = 2.29%). These summary statistics are similar to those reported by previous authors. Negative event day returns are particularly rare in the last years of our sample, presumably reflecting the effect of larger demand shocks by growing index funds.

B. Variables

The measurable right-hand-side variables in equation (5) are arbitrage risk *A* and demand shock size *S*. We first discuss arbitrage risk. Definition (2) implies that arbitrageurs consider all available assets as potential substitutes. However, solving for optimal zero-net-investment portfolio weights on each of thousands of assets for each of 259 stocks is too difficult both for us and for D. E. Shaw. Indeed, the transaction costs of hedging with many distinct assets would be prohibitively high. The first task is to select an appropriately restricted set of potential substitutes. To avoid focusing on the wrong substitutes, we construct both a simple and more detailed measure under alternative assumptions about the universe of potential substitutes considered by arbitrageurs. The substitutes we consider are (1) the market portfolio and (2) three stocks that match the subject stock on industry and as closely as possible on size and book-to-

8. More precisely, index change announcements are made after the market close, so the event day is the next trading day.

9. Marion Laboratories, added May 4, 1988, fell over 13.4% on the event day amid the announcement that a rival drug company was going to produce a generic version of its biggest-selling product. The next poorest event-day performance out of 259 other events was a drop of 2.5%. We exclude Marion Laboratories to prevent it from dominating the regression results. Median regressions that include this data point yield virtually identical results to the OLS regressions we report.

market. We denote the resulting arbitrage risk measures A_1 and A_2, respectively.[10]

To illustrate our procedure, we describe how we estimate A_2 for Church's Fried Chicken. Church's was added to the S&P 500 Index on November 16, 1977, and deleted March 1, 1989, following its acquisition. We place all Center for Research in Security Prices (CRSP) stocks in Church's Fama and French (1997) industry as of November 1977 into quintiles by the absolute value of the difference between their market capitalizations and Church's market capitalization and also by the difference between their book-to-market ratios and Church's book-to-market ratio. We sort these stocks by market capitalization difference quintile and then by book-to-market difference quintile. We then select the three stocks at the top of this sort. To break ties among stocks with the same values for both quintiles, we select the one with the smallest absolute difference in capitalization. This procedure selects Wendy's, Morrison Restaurants (now Ruby Tuesday), and Friendly's as industry-, size-, and book-to-market-matched substitutes for Church's Fried Chicken in November 1977.

The final step is to determine the variance of the zero-net-investment portfolio long \$1 in Church's and short a net of \$1 across these three substitute stocks, where the short position is set to minimize the variance of the overall position. We could regress Church's historical returns on those of the three substitutes, constrain the sum of the three slope coefficients to equal unity, and call the residual variance $A_{2\text{Church's}}$. This is essentially what we do, except we consider excess returns. We specify Church's excess returns as a combination of the excess returns of Church's substitutes. The benefit of this is that we need not constrain the sum of the slope coefficients in order to maintain a zero-net-investment. To illustrate, consider our final model for Church's excess returns, estimated using daily data over the calendar day $[-365, -20]$ preevent window:

$$\begin{aligned}(R_{\text{Church's},\,t} - R_{ft}) = {} & .108(R_{\text{Wendy's},\,t} - R_{ft}) \\ & + .259(R_{\text{Morrison},\,t} - R_{ft}) + .109(R_{\text{Friendly's},\,t} - R_{ft}).\end{aligned}$$

The implied arbitrage strategy is for every \$1 long in Church's, to short \$1 in T-bills, short \$.108 in Wendy's, long \$.108 more in T-bills, short \$.259 in Morrison, long \$.259 more in T-bills, short \$.109 in Friendly's, and long \$.109 more in T-bills. Then $A_{2\text{Church's}}$ is the residual variance of this portfolio.[11] We

10. Why do we not consider hedging with options? It is possible to hedge the short stock position perfectly with a combination of a put, a call, and a bond. However, if put-call parity holds, the stock and the options will be similarly mispriced, and one cannot profit from the fundamental mispricing. If put-call parity does not hold, there is a much larger inefficiency at hand. Put differently, hedging a long position in a stock using the stock's own options is akin to being simultaneously long and short the same share. The position that results is indeed riskless, but it also contains no profit opportunity.

11. The model does not identify the arbitrageur's investment horizon. Our procedure implicitly uses a daily measure, which is surely much too short. In general, however, stocks that are hard to hedge at daily frequencies are also hard to hedge at longer frequencies, so the cross-sectional ranking is preserved.

TABLE 1 The Arbitrage Risk of Stocks Added to the S&P 500 Index

	Mean	SD	Minimum	Median	Maximum
A_{1i}	.000372	.000255	.000033	.000311	.002219
A_{2i}	.000409	.000275	.000034	.000334	.002277
$\mathrm{Var}(R_{it})$	.000469	.000318	.000039	.000379	.002374
$\mathrm{Var}(R_{it} - R_{ft})$	.000468	.000314	.000039	.000379	.002375
E_{1i}	.000098	.000112	−.000006	.000050	.000791
E_{2i}	.000061	.000089	−.000004	.000027	.000726
$E_{1i}/\mathrm{Var}(R_{it} - R_{ft})$	.183	.116	−.011	.151	.585
$E_{2i}/\mathrm{Var}(R_{it} - R_{ft})$	.109	.103	−.016	.078	.471

NOTE.—The sample includes 259 stocks added to the S&P 500 Index between September 1976 and September 1989. We construct two stock-specific arbitrage-risk measures. Both models were estimated with daily returns over a [−365, −20] calendar day window. Arbitrage-risk measure A_{1i} is the variance of the residuals of the model $R_{it} - R_{ft} = \beta_{1i}(R_{mt} - R_{ft})$, where R_{mt} is the return on CRSP's value-weighted AMEX/NASDAQ/NYSE index and R_{ft} is the T-bill return. Daily values for R_{ft} were approximated with appropriate roots of the corresponding monthly return. Measure A_{2i} is the residual variance of $R_{it} - R_{ft} = \beta_{1i}(R_{sub1it} - R_{ft}) + \beta_{2i}(R_{sub2it} - R_{ft}) + \beta_{3i}(R_{sub3it} - R_{ft})$, where R_{sub1it}, R_{sub2it}, and R_{sub3it} denote returns on three industry-, size-, and book-to-market-matched "substitute" stocks. We use the Fama and French (1997) industry classifications to match on industry. See text for details of the substitute selection procedure. $\mathrm{Var}(R_{it})$ and $\mathrm{Var}(R_{it} - R_{ft})$ are estimated over the same interval as the arbitrage-risk measures. E_{1i} is the "explained" variance $\mathrm{Var}(R_{it} - R_{ft}) - A_{1i}$. E_{2i} is defined analogously.

compute A_1 and A_2 similarly for each of the 259 stocks in the sample. This procedure is very similar to that used by Pontiff (1996). Pontiff selects 10 open-end funds as the universe of potential substitutes for the closed-end funds in his sample.

Table 1 reviews the estimation procedures and gives summary statistics. For reference, we compute summary statistics for the total variance of added stocks' excess returns. We also compute the absolute amount and the fraction of total variance explained by substitutes' returns.[12]

The general message in table 1 is that individual stocks do not have perfect substitutes. Risk measures A_1 and A_2 are similar in magnitude, highly correlated (about .98), and on the same order of magnitude as the risk faced by not hedging at all [$\mathrm{Var}(R_{it})$]. The last four rows of the table indicate that good substitutes can reduce the variance of speculative returns by one-fifth for typical stocks and by up to three-fifths for a very few stocks. But for some stocks, hedging is essentially impossible. They have no decent substitutes. Short-sellers of these stocks may as well hold a risk-free asset rather than search for substitutes. These results agree with Roll (1988): for daily returns, the average R^2 is around .20 for his best-fitting models.

Stocks with good substitutes include American Barrick Resources (substitutes include other gold mining companies) and Standard Oil (other oil companies). In general, small stocks tend to have higher arbitrage risk than large stocks. In our sample, the correlation between log (market value) and A is negative and very highly significant.

Index fund demand shock size, S, is easier to measure. Since the S&P 500

12. The zero-net-investment portfolio models do not include a constant (it has no interpretation), so it is not identically true that total variance can be decomposed into residual variance (arbitrage risk) and explained variance. This accounts for the negative values we occasionally obtain for the latter.

TABLE 2 **The Demand Shock to Stocks Added to the S&P 500 Index (Selected Years)**

	1976	1980	1984	1988	1992	1996
S&P 500–tracking index fund capitalization ($ billion)*	19	35	68	135	255	475
S&P 500 total capitalization ($ billion)	662	926	1,217	1,897	3,015	5,626
Size of demand shock as:						
% of market capitalization	2.9	3.8	5.6	7.1	8.5	8.4
Dollar value (median; $ million)†	11	31	42	96	183	418

NOTE.—All values are end-of-year, and all monetary values are in current dollars. S&P 500–tracking index fund capitalization is an estimate of the aggregate capitalization of all public and private funds indexed to the S&P 500. S&P 500 total capitalization is the aggregate market capitalization of all 500 firms in the index at the end of the given year. Data for these two items were provided by Standard and Poor's Index Services. The percentage of a newly added firm's capitalization demanded by index funds is the ratio of S&P 500 index fund capitalization to S&P 500 total capitalization. The reported dollar value demand shock is the median shock over all firms in our sample that were added that year.

* Standard and Poor's had estimates available only for years 1985 through 1997. The values for earlier years were estimated by fitting an exponential curve through the 1985–1997 data, which fits with an R^2 of .97.

† Due to data availability, the 1976 value applies to firms added to the index between September 22, 1976, and December 31, 1976, and the 1996 value applies to firms added between January 1, 1996, and May 21, 1996.

is a value-weighted index, the ratio of total S&P 500-indexed assets to total S&P 500 capitalization identifies the percentage of each included firm that is demanded by index funds. Table 2 summarizes the growth of index funds over the period 1976–96, based on estimates provided by Standard and Poor's Index Services. By these estimates, index funds grew 2,400% between 1976 and 1996, increasing from $19 billion to $475 billion. Over the same period, the total capitalization of the S&P 500 grew by 750%, from $662 billion to $5,626 billion. Thus the percentage of an added firm's capitalization demanded by index funds has grown strongly, from about 3% in 1976 to more than 8% in 1996.[13]

The fourth row of table 2 calculates the size of the demand shock, in current millions of dollars, for the year's median-size addition. It is simply the median dollar value of the percentage shocks reported in row 3. This variable varies across firms within a year, in contrast to the percentage shock variable, which varies only across years.[14] For purposes of comparison, we gather data from 1992 and 1996 for this row.

13. The results of Blume and Edelen (2001) suggest that a better measure of the effective demand shock is the index demand relative to the public float, not relative to the total shares outstanding. Unfortunately, we are not aware of an electronic source for insider holdings that covers our sample period.

14. In practice, some index funds may wait for a few days or even until the end of the month to rebalance. The validity of our assumption—that the full shock is satisfied immediately—requires that arbitrageurs be able to calculate the demand they will ultimately need to satisfy. Competition among these arbitrageurs ensures that the impact of the full demand shock will be reflected at announcement.

C. Empirical Results

Equation (5) predicts that event returns should be increasing in arbitrage risk and shock size. Table 3 reports ordinary least squares (OLS) estimates of models of announcement day returns. Specifications 1 and 2 show that arbitrage risk and demand shock size are both positively and significantly related to event returns. Given their very high correlation with each other (above .98), it is not surprising that the arbitrage risk measures A_1 and A_2 have similar explanatory power. Specification 3 illustrates that the demand shock measured in percentage terms is also strongly related to event returns. These results broadly confirm two qualitative predictions of our model. Announcement day returns are significantly higher for stocks without close substitutes and for stocks hit by large index fund demand shocks.

The model also predicts that event day returns ought not respond to the component of risk that can be hedged away, only to the arbitrage risk component. Specification 4 tests this prediction. The explained component of risk is not significantly related to event returns. These results are similar to results in Pontiff (1996), in which the cross section of average closed-end fund discounts depends only on the unhedgeable component of closed-end fund risk.

Specification 5 considers the possibility that the arbitrage risk coefficient is biased because it is correlated with the omitted variable h. One could imagine an extreme inefficient market model in which arbitrage is completely unimportant, so the aggregate demand curve is determined entirely by the heterogeneity of nonarbitrageurs' beliefs. To test this interpretation, we proxied for the heterogeneity of investor beliefs using the heterogeneity of analysts' earnings forecasts from I/B/E/S. We measured *analyst dispersion* as the standard deviation of analysts' forecasts for one-year-ahead earnings per share (EPS), divided by the mean estimate. We exclude stocks that have fewer than five such estimates as of the event date. Specification 5 reports that this proxy for h is not strongly related to event returns, and controlling for it does not diminish the coefficient on arbitrage risk. However, consistent with this hypothesis's premise that arbitrage risk is correlated with h, we do find a significant positive correlation of .138 between arbitrage risk and the heterogeneity of earnings forecasts. We find similar results for a specification that measures *analyst dispersion* as the standard deviation of analysts' forecasts for one-year-ahead EPS, divided by the stock price (not reported). It seems likely that the weak effect of *analyst dispersion* is due to its theoretically ambiguous relationship to true heterogeneity h. That is, one can imagine that career concerns cause analysts to give similar forecasts for precisely those stocks in which true heterogeneity and uncertainty is high. The results of Bagwell (1992) and Kandel et al. (1999) demonstrate that heterogeneity does affect the slopes of demand curves in situations where heterogeneity can be measured.

Equation (5) of the model predicts that arbitrage risk and shock size do

TABLE 3 **The Effects of Arbitrage Risk and Demand Shock Size on S&P 500 Addition Event Returns**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	.023**	.023**	−.002	−.002	.002	.027**	.023**	.024**
	(.003)	(.003)	(.006)	(.006)	(.007)	(.002)	(.002)	(.003)
A_{1i}	16.234**		16.294**	14.131*	16.746*			
	(5.953)		(5.834)	(6.518)	(7.038)			
A_{2i}		14.257*						
		(5.761)						
E_{1i}				11.395				
				(16.360)				
Analyst dispersion$_i$					−.006			
					(.015)			
Shock mil$_i$	.000043*	.000043*						
	(.000021)	(.000021)						
Shock pct$_i$			.00509**	.00496**	.00428**			
			(.00105)	(.00103)	(.00126)			
A_{1i} × Shock mil$_i$						.206**		
						(.070)		
A_{1i} × Shock pct$_i$							4.081**	3.901**
							(1.112)	(1.320)
Analyst dispersion$_i$ × Shock pct$_i$								−.00056
								(.00060)
$\bar{R}^2$	.039	.036	.102	.104	.080	.037	.061	.060
N	259	259	259	259	177	259	259	177

NOTE.—This table presents OLS models of S&P 500 Index addition announcement day returns. The dependent variable is the excess return over the market ($R_{it} - R_{mt}$). The sample includes 259 stocks added to the S&P 500 between September 1976 and September 1989, or as limited by data. The construction and summary statistics of arbitrage risk A and explained risk E are summarized in table 1. Analyst dispersion$_i$ is the standard deviation of analysts' forecasts for 1-year-ahead earnings per share, divided by the mean estimate, for stocks that have such estimates outstanding by at least five analysts at the time of the addition (mean = .094, median = .060, SD = .118, $N = 177$). Shock mil$_i$ is the dollar value of S&P 500 index fund demand, in millions of current dollars (sample mean = \$64 [million], median = \$51 [million], SD = \$64 [million], $N = 259$). Shock pct$_i$ is the percentage of the added firm's capitalization demanded by index funds (sample mean = 5.0, median = 5.4, SD = 1.3, $N = 259$). Heteroskedasticity-robust standard errors are in parentheses.

* Significant at the 5% level.

** Significant at the 1% level.

not enter linearly. In particular, when the demands of nonarbitrageurs are relatively inelastic (h is high), nonarbitrageurs do not supply much stock to index funds. Only arbitrageurs will be willing to satisfy the demand shock. In this case, the interaction of arbitrage risk and shock size is what best explains event returns. Specifications 6, 7, and 8 confirm that the interaction of arbitrage risk and shock size is a strong explanatory variable of event returns.

One potential challenge to our interpretation of these results is an imperfect indexing effect: if indexers minimize tracking error but face transaction costs, then all else (such as shock size) equal, they will prefer to add the stocks that are less correlated with other securities. In this case arbitrage risk could pick up some of the unmeasured component of actual shock size. From what evidence we could gather, this does not appear to be a compelling hypothesis. Announcement day volume, relative to historical average volume or as a percent of dollar shock size, is not differentially higher for high-arbitrage-risk stocks (not reported). In addition, it seems possible, if not very likely, that the loss of fund inflows due to increased tracking error far outweighs the transaction costs saved by imperfect indexing, so indexing funds—at least those that solicit outside investors—have incentives to index as perfectly as they can.

Taken as a whole, the results indicate that arbitrage fails to flatten demand curves for stocks and that this happens in part because perfect substitutes are not available. On the other hand, we note that several specifications include a constant term of 2%. In addition, the explanatory power of our regressions is not high, though this is not unusual for studies that try to explain a cross section of daily returns. It is unclear whether these shortcomings reflect the fact that our regression model is misspecified (vs. eq. [5]), that we are ignoring other important effects, or that the noise component of event day returns dominates the signal and even a perfect structural model would be unable to explain much more of the sample variation.

We perform two additional exercises to shed light on the economic significance of the results. First, we plot mean cumulative abnormal returns for added stocks to examine whether the event day effects are quickly reverting or not. This addresses the question of whether our results reflect temporary price pressures or longer-run downward-sloping demand curves. Second, in the next subsection we discuss the estimates of slopes of demand curves implicit in our results and compare them with estimates from other sources.

Shleifer (1986), Dhillon and Johnson (1991), Beneish and Whaley (1996), and Lynch and Mendenhall (1997) argue that the addition effect is at least partially permanent, while Harris and Gurel (1986) use a different methodology under which the effect completely and quickly reverses. Our model predicts that there will be at least some reverting component, to reward arbitrageurs for accommodating the index funds' demands by selling shares to them. All prior research points to this conclusion.

Unfortunately, long-run inference is notoriously difficult. The signal component of the average 3% abnormal return quickly becomes swamped by the noise added by new fundamental news within a few days after the addition. To minimize the influence of extraneous factors on cumulative abnormal returns and to perform inference as precisely as possible, we have cleaned the sample by eliminating 68 addition events that occurred within 2 days of reports of other fundamental news reported in NEXIS (which subsumes the *Wall Street Journal Index*, for instance).[15]

Figure 1 plots mean cumulative abnormal returns for added stocks by shock size. To control for the interactive effect with arbitrage risk, we exclude stocks in the highest and lowest quartiles of the arbitrage-risk distribution from the plot. Figure 1 confirms that larger shocks induce larger price responses, and the average does not appear to reverse within 20 trading days. Given the standard errors reported in the table, the two series are statistically different on the announcement day (as is also clear from table 3), but the rapid growth in the variance of cumulative returns causes the series to be statistically indistinguishable within 10 days. Indeed, after a few weeks, one cannot reject the hypothesis that the effect has completely reversed.

Figure 2 plots mean cumulative abnormal returns for the added stocks, by arbitrage risk. To control for the interactive effect with shock size, we exclude stocks in the highest and lowest quartiles of the shock size distribution. Figure 2 confirms that high-arbitrage-risk stocks experience larger average price responses. These abnormal returns are not quickly mean reverting. As in figure 1, the mean returns are statistically different on the announcement day but quickly become statistically indistinguishable from each other and from zero. Similar figures result from using other arbitrage risk measures or from plotting medians instead of means.[16]

Beyond 20 days, the direction that the mean cumulative abnormal return estimate takes becomes more sensitive to methodology. Under most risk adjustments that we have tried, we find a partial reversion within several months, which is consistent with prior literature. We do not report these estimates since at least five other papers report similar results for essentially the same sample.

It is important that a partial reversion is as consistent with the model as a full reversion—either way arbitrageurs on average make a profit. Unfortunately, the model cannot specify the horizon over which the reversion should occur. Harris and Gurel suggest that full reversion takes place within 2 weeks, while DeBondt and Thaler (1985) look for mean reversion over horizons of

15. Confounding news events include 37 reports about significant changes in investment practices and prospects of the firm or its industry, including acquisitions, capital expenditures, new product announcements, and price changes; 18 earnings announcements and earnings forecasts; six capital structure news reports, including bond rating changes, dividend changes, and security issues; six litigation news reports; and one report of executive turnover.

16. Figure 1 looks the same using the full, unclean sample. Figure 2 looks similar but, apparently due to noise, the cumulative abnormal return (CAR) of the low-arbitrage-risk series jumps up at day $+5$, while at the same time the CAR of the high-arbitrage-risk series jumps down, causing the two plots to cross. Thus we caution the reader against depending too much on the plots beyond a very few days after the event.

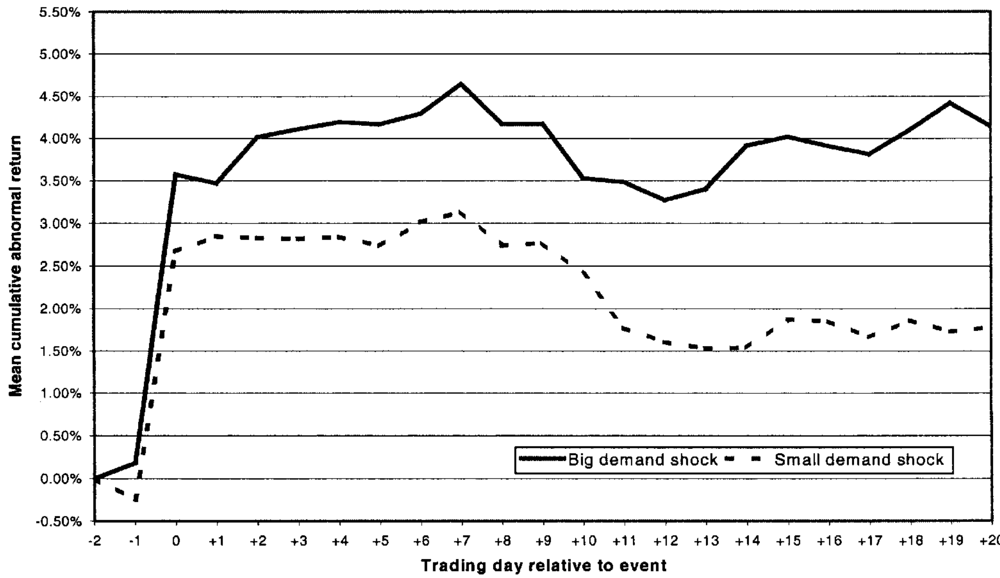

FIG. 1.—Mean cumulative abnormal returns for stocks added to the S&P 500, by size of index fund demand shock. The sample includes 191 stocks that were added to the S&P 500 between September 1976 and September 1989 and were not the subject of contemporaneously reported news. To control for the level of arbitrage risk, we exclude stocks in the extreme two quartiles of the arbitrage-risk distribution (measure A_1). We split the remaining 96 stocks into above-median and below-median demand shock groups. Cumulative abnormal returns are calculated by summing returns over the CRSP value-weighted market index. The standard error of the estimates in each series is approximately .40% at day 0, .90% at day +10, and 1.10% at day +20.

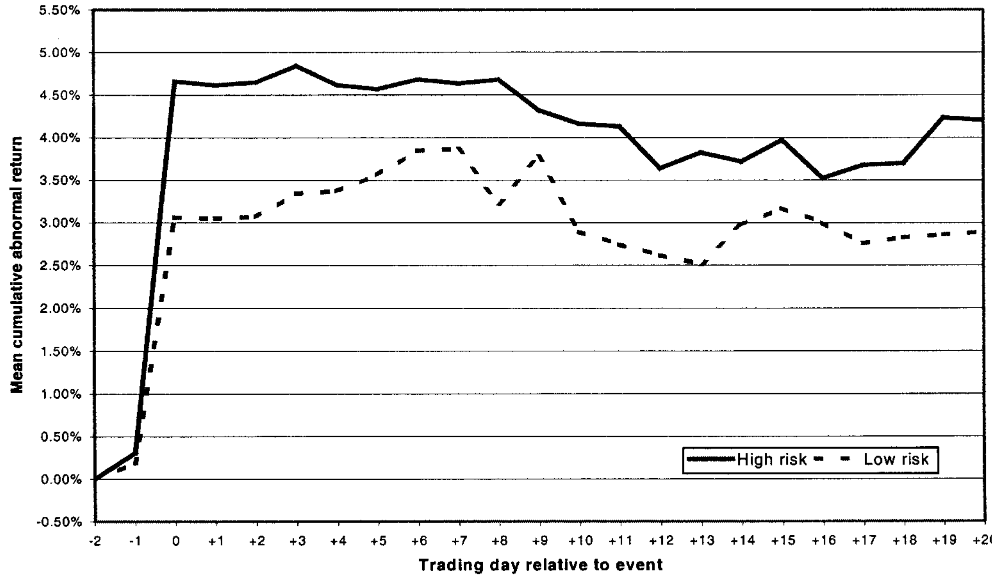

FIG. 2.—Mean cumulative abnormal returns for stocks added to the S&P 500, by level of arbitrage risk. The sample includes 191 stocks that were added to the S&P 500 between September 1976 and September 1989 and were not the subject of contemporaneously reported news. To control for the size of the index fund demand shock, we exclude stocks in the extreme two quartiles of the shock size distribution. We split the remaining 96 stocks into above-median and below-median arbitrage-risk groups, using measure A_1. Cumulative abnormal returns are calculated by summing returns over the CRSP value-weighted market index. The standard error of the estimates in each series is approximately .40% at day 0, .80% at day +10, and 1.00% at day +20.

several years. We conclude that some reversion exists, but it takes longer than the 10 days or so in which reliable inference in this sample is still possible. In a sense, the fact that reversion is difficult to document precisely is consistent with the model. If reversion were too quick and dependable, arbitrage would be riskless and unlimited, and in equilibrium there would be no price effect in the first place.

D. Elasticity Estimates

In this subsection, we discuss the slopes of aggregate demand curves implicit in our regression results. The convention in prior literature is to focus on the price elasticity of demand, that is, the percent change in demand per percent increase in price. The model in this article, on the other hand, implies that demand curves have constant semi-elasticities, that is, constant dollar changes in demand per percent increases in price. Although we see no theoretical reason why demand curves for stocks should be isoelastic, we will assume that they are for the sake of comparing our results with prior estimates.

To be clear, the elasticity measure we are interested in is $(\%\Delta x/\%\Delta P)$, where x can be taken as excess demand since supply curves for stocks are vertical. Conveniently, this means the price elasticity of demand can be identified from supply shocks—changes in total shares outstanding—as well as demand shocks. In some cases, such as the index additions, it is easier to think about the demand shock as a reduction in public float.

The perfect market benchmark is, of course, that the elasticity is negative infinity. This is the perfectly flat demand curve. If a stock with an elasticity of negative infinity is even a little overpriced (underpriced), investors will want to sell short (buy) an infinite amount of it. The closest estimate to negative infinity that we find is also the first estimate, made by Scholes (1972). Scholes regresses the abnormal return at the announcement of a secondary distribution of a large block on the size of the large block as a percent of the firm. Assuming that any information associated with such an announcement is absorbed into the constant term, the inverse of the coefficient on percent block size is the elasticity of interest. This coefficient is close to zero (the regression itself has an R^2 of .0004), leading Scholes to estimate an elasticity of $-3{,}000$. He concludes that demand curves for stocks are essentially flat. Table 4 summarizes this and other elasticity estimates for individual stocks.[17] The estimates are not always easy to find in the original papers (they are not always the focus), so we also indicate the appropriate page number.

More recent elasticity estimates suggest that demand curves are much steeper than Scholes's estimate. Shleifer (1986) notes that the typical S&P 500 addition effect in his sample is around 3%. He also points out that, at the time of the sample, index funds demanded about 3% of the newly added stocks. If one attributes the entire addition effect to a shift in uninformed

17. Gompers and Metrick (2001) estimate the price elasticity of demand for the market portfolio as a whole, using estimates of the elasticity of individual stocks as a crucial input.

TABLE 4 Estimates of the Price Elasticity of Demand for Individual Stocks

Source	Elasticity Estimate	Sample
Scholes (1972, p. 195)	−3,000	345 secondary distributions of large blocks on NYSE between 1961 and 1965
Shleifer (1986, inferred)	−1	246 additions to the S&P 500 between 1976 and 1983
Loderer, Cooney, and Van Drunen (1991, p. 640)	−11.12 (mean) −4.31 (median)	409 announcements of equity issues by regulated firms between 1969 and 1982
Bagwell (1992, p. 97)	−1.65	31 Dutch auction repurchases between 1981 and 1988
Kandel, Sarig, and Wohl (1999, p. 235)	−37.2 (mean) −21.0 (median)	27 Israeli IPO auctions between 1993 and 1996
Kaul, Mehrotra, and Morck (2000, p. 911)	−10.5	292 stocks affected by Toronto Stock Exchange 300 index weights change effective November 1996
The current article (specification 8, table 3)	−11.72 (at twenty-fifth percentile arbitrage risk) −8.24 (at median arbitrage risk) −5.57 (at seventy-fifth percentile arbitrage risk)	177 additions to the S&P 500 between 1976 and 1989

NOTE.—This table reports estimates of the price elasticity of demand for individual stocks. The measure is the percent change in demand associated with a 1% higher price. The perfect market benchmark is negative infinity. In cases where the author uses a different definition, we transform the results to obtain comparable estimates. The estimates are ordered by publication date. The page on which the relevant data is published is reported.

demand, as opposed to an information or liquidity change, these figures imply an elasticity estimate of about −1. The results of Harris and Gurel (1986) imply a similar estimate, but they put less emphasis on an explanation based on sloping demand curves.

Loderer et al. (1991) estimate the price elasticity of demand from a sample of announcements of primary offerings by regulated firms. These announcement effects, the authors argue, are unlikely to be driven by adverse information. Dividing the percent increase in supply due to the new shares by the announcement effect (which is typically negative), they estimate a mean price elasticity of demand of −11.12 and a median elasticity of −4.31. These estimates, like Shleifer's, attribute the entire price effect to the change in excess demand. In additional analysis, Loderer et al. attempt to control for the impact of adverse information directly and explore the cross section of implied elasticities in more detail.

Bagwell (1992) studies a small sample of firms that conducted Dutch auction share repurchases. She expresses her results from the perspective of the repurchasing firms, which essentially face upward-sloping supply curves for repurchased shares. This is the reflection of downward-sloping demand curves from the investors' perspective. She finds that to repurchase 15% of outstanding stock in a Dutch auction, the median firm must offer a 9.1% premium. This leads to a demand elasticity estimate of -1.65.

Kandel et al. (1999) study a sample of 27 Israeli IPOs that are conducted as nondiscriminatory (uniform price) auctions. Their data allow them to completely trace out the demand schedule for these stocks. Around the auction's clearing price, they estimate the median firm's elasticity as -21.0 and the mean firm's elasticity as -37.2. As it is not clear that Scholes-type arbitrageurs would be able to enter bids in this auction, the relatively flat demand curve estimates from this sample are somewhat more surprising.

Kaul et al. (2000) investigate the price effects of a definitional change in the Toronto Stock Exchange 300 (TSE) index. This change produced changes in demands by index funds that track the TSE. Similar to Scholes (above) and us (below), they estimate an elasticity from the inverse of the slope coefficient in regressions of individual securities' price responses on the implied percent changes in public float. Their elasticity estimate is -10.5. It is interesting that, in this experiment, the effects of the reweighting are manifest primarily at the effective date, not the announcement date, and this is the point at which Kaul et al. take the price response to estimate the elasticity. The early announcement creates the opportunity to front-run (as in Beneish and Whaley 1996). The fact that there was no run-up in price that eliminated the effective day abnormal return (at which time front-running arbitrageurs would distribute their accumulated supply) suggests that "arbitrageurs" simply did not notice that this change was going to happen. Perhaps they did not pick up on the change until the effective date, and perhaps they never did. In that case, the estimated elasticity reflects mainly the heterogeneity of nonarbitrageurs' beliefs about the value of these stocks. Another possibility is that the precise basis for the weights change was not known ex ante, leaving arbitrageurs to guess about the exact size of the demand shock. However, they still had the opportunity to diversify across stocks subject to weights changes and thus they could diversify away some of this uncertainty. In any case, the key point is that the Scholes arbitrage was clearly inoperational.

We also report the elasticity estimates that are implied on our own results. The message of this article is that the slopes of demand curves for stocks depend on arbitrage risk. We, therefore, compute the implied elasticity for stocks at the twenty-fifth percentile, at the median, and at the seventy-fifth percentile of arbitrage risk using specification 8 of table 3. A stock at the median of arbitrage risk has an elasticity of -8.24. The calculation for this is $-8.24 = 1/(3.901 * .000311 * 100 * -1)$, where 3.901 is the coefficient on the interaction of arbitrage risk and percent shock in specification 8, .000311 is the median stock's arbitrage risk from table 2, 100 is a scale factor to turn decimal returns to percentage changes, and -1 is the percent decrease in demand by arbitrageurs and nonarbitrageurs that accommodates the percent demand shock. A stock at the twenty-fifth percentile of arbitrage risk has a significantly flatter demand curve with an elasticity of -11.72. A stock that is at the seventy-fifth percentile of arbitrage risk has a significantly steeper demand curve with an elasticity of -5.57. In other words, a given percentage demand shock causes twice the response in a stock at the seventy-fifth percentile of arbitrage risk as compared with a stock at the twenty-fifth percentile.

The difference between our estimates and Shleifer's estimates, which also come from the S&P 500 setting, is mainly due to the constant term in specification 8. This term may reflect an information or liquidity component to the addition effect, in which case our estimate is more accurate than Shleifer's. On the other hand, our specification may be incorrect, and the entire addition effect may reflect downward-sloping demand curves, in which case Shleifer's estimate is preferable. Some support for our approach comes from the fact that our estimates are somewhat less extreme than his relative to estimates from other settings. Whether the elasticity is -1 or -8.24, though, the essential theoretical conclusion is the same: arbitrage fails to flatten demand curves for stocks.

IV. Summary and Implications

Scholes (1972) suggests that arbitrage between perfect substitutes keeps the demand curves for stocks flat. While between-stock arbitrage is widely pursued by real "arbitrageurs," most stocks do not have very close substitutes, and such strategies are far from riskless. This article presents empirical evidence from S&P 500 Index additions that suggests that arbitrage risk is an important determinant of the demand curve's slope. We believe the results are most consistent with the interpretation that risk discourages arbitrage from flattening demand curves for stocks.

The results of this article lead us to differ with Ross (1987), who writes, "not to say that the intuition and the theories of finance cannot be fit into the framework of supply and demand, rather that doing so does not gain us much. The fit is awkward and irrelevant at best" (p. 30). A methodical investigation of the limits of arbitrage, and the implications of these limits for the shapes of excess demand curves for stocks, seems likely to improve our understanding of the growing set of empirical findings that are difficult to explain with models that assume unlimited arbitrage.

Further empirical research on how arbitrage risk affects prices could proceed along two broad lines. The first line would attempt to confirm or refute our proposition that the response to excess demand shocks depends on arbitrage risk. Potential shocks may include equity issues, repurchases, or takeovers. A stumbling block for studies like these, however, is the difficulty of controlling for new information associated with the shock. Experiments as informationally clean as S&P 500 Index changes are hard to find.

A second line of research would explore the impact of arbitrage that is limited by arbitrage risk. The guiding hypothesis is that true anomalies are likely to be severest in settings where arbitrage is most difficult. This observation has been made previously by De Long et al. (1990), Pontiff (1996), and Shleifer and Vishny (1997), among others. To this end, we find it particularly interesting that arbitrage risk is higher for smaller stocks. This may help to explain Loughran and Ritter's (2000, p. 363) observation that "just about every known stock market pattern is stronger for small firms than for large firms."

An important observation about the implications of idiosyncratic arbitrage risk for various market anomalies has been made by Campbell (2000) and Shleifer (2000) in their discussions of the current article. By its nature, arbitrage risk is, in principle, diversifiable. But one simply cannot form a portfolio of S&P 500 additions, because there is only one change every few weeks. Similarly, it is not easy to form a portfolio of initial public offerings at the end of the first trading day (at which point several authors have argued that they are overpriced, on average). In general, the more mispricings that can be identified and traded upon at once, the less influence arbitrage risk should have on any particular mispricing, and therefore the smaller the average mispricing should be in equilibrium. Put differently, the abnormal returns on a given portfolio of events should be smaller and abnormal returns on the individual events should be less related to arbitrage risk in periods when there are more events in the portfolio. The results of Baker and Savasoglu (2002) generally support these predictions. They find that excess returns to individual risk-arbitrage deals are increasing in measures of the risk that the merger deal goes through, but decreasing in the total number of merger deals outstanding at the time.

As a final remark, the behavioral asset pricing models that are emerging must always explain first why arbitrage fails before they explore how irrational trading affects prices. This point is emphasized by Shleifer (2000). Our results suggest that risk resulting from a lack of substitutes may itself be enough to justify an assumption of limited arbitrage. Arbitrage risk can be added to a list of factors that may inhibit market efficiency: informational and transaction costs (Stoll 2000), heterogeneous beliefs about fundamental value (Bagwell 1992), noise trader sentiment risk (De Long et al. 1990), short-sales constraints (Chen et al. 2002), and agency costs in delegated arbitrage (Shleifer and Vishny 1997). In contrast to most other limits on arbitrage, however, arbitrage risk is relatively easy to measure.

References

Asquith, P., and D. W. Mullins. 1986. Equity issues and offering dilution. *Journal of Financial Economics* 15 (January/February): 61–89.

Bagwell, L. S. 1992. Dutch auction repurchases: An analysis of shareholder heterogeneity. *Journal of Finance* 47 (March): 71–105.

Baker, M., and Savasoglu, S. 2002. Limited arbitrage in mergers and acquisitions. *Journal of Financial Economics* 64 (April): 91–116.

Beneish, M. D., and Whaley, R. E. 1996. An anatomy of the "S&P game": The effects of changing the rules. *Journal of Finance* 51 (December): 1909–30.

Blume, M. E., and Edelen, R. M. 2001. The Yahoo! squeeze: Index tracking institutions and restricted float. Wharton School working paper. Philadelphia: University of Pennsylvania

Bradley, M.; Desai, A.; and Kim, E. H. 1988. Synergistic gains from corporate acquisitions and

their division between the stockholders of target and acquiring firms. *Journal of Financial Economics* 21 (May): 4–40.
Campbell, J. Y. 2000. Asset pricing at the millennium. *Journal of Finance* 55 (August): 1515–67.
Chen, J.; Hong, H.; and Stein, J. C. 2002. Breadth of ownership and stock returns. *Journal of Financial Economics* (forthcoming).
DeBondt, W. F. M., and Thaler, R. 1985. Does the stock market overreact? *Journal of Finance* 42 (July): 557–81.
De Long, J. B.; Shleifer, A.; Summers, L.; and Waldmann, R. 1990. Noise trader risk in financial markets. *Journal of Political Economy* 98 (August): 703–38.
Dhillon, U., and Johnson, H. 1991. Changes in the Standard and Poor's 500 list. *Journal of Business* 64 (January): 75–85.
Fama, E. F., and French, K. R. 1997. Industry costs of equity. *Journal of Financial Economics* 43 (February): 153–93.
Gatev, E.; Goetzmann, W. N.; and Rouwenhorst, K. G. 2000. Pairs trading: Performance of a relative value arbitrage rule. Yale School of Management working paper. New Haven, Conn.: Yale University.
Goetzmann, W. N., and Garry, M. 1986. Does delisting from the S&P 500 affect stock price? *Financial Analysts Journal* 42 (March/April): 64–69.
Gompers, P. A., and Metrick, A. 2001. Institutional investors and equity prices. *Quarterly Journal of Economics* 116 (February): 229–59.
Green, R. C., and Rydqvist, K. 1997. The valuation of nonsystematic risks and the pricing of Swedish lottery bonds. *Review of Financial Studies* 10 (Summer): 447–80.
Grossman, S. J., and Miller, M. H. 1988. Liquidity and market structure. *Journal of Finance* 43 (July): 617–33.
Harris, L., and Gurel, E. 1986. Price and volume effects associated with changes in the S&P 500: New evidence for the existence of price pressures. *Journal of Finance* 41 (September): 815–29.
Holthausen, R.; Leftwich, R.; and Mayers, D. 1990. Large-block transactions, the speed of response, and temporary and permanent stock-price effects. *Journal of Financial Economics* 26 (July): 71–95.
Ip, G. 1998. "S&P effect" often ignites some stocks as they are being added to 500 Index. *Wall Street Journal* (January 30).
Jensen, M. C. 1994. Arbitrage, information theft, and insider trading. In P. Newman, M. Milgate, and J. Eatwell (eds.), *The New Palgrave Dictionary of Money and Finance*, pp. 50–52. Vol. 1. New York: Macmillan.
Kandel, S.; Sarig, O.; and Wohl, A. 1999. The demand for stocks: An analysis of IPO auctions. *Review of Financial Studies* 12 (Summer): 227–47.
Kaul, A.; Mehrotra, V.; and Morck, R. 2000. Demand curves for stocks *do* slope down: New evidence from an index weights adjustment. *Journal of Finance* 55 (April): 893–912.
Loderer, C.; Cooney, J.; and Van Drunen, L. 1991. The price elasticity of demand for common stock. *Journal of Finance* 46 (June): 621–51.
Loughran, T., and Ritter, J. R. 2000. Uniformly least powerful tests of market efficiency. *Journal of Financial Economics* 55 (March): 361–90.
Lynch, A. W., and Mendenhall, R. R. 1997. New evidence on stock price effects associated with changes in the S&P 500 Index. *Journal of Business* 70 (July): 351–83.
Pontiff, J. 1996. Costly arbitrage: Evidence from closed-end funds. *Quarterly Journal of Economics* 111 (November): 1135–52.
Richards, A. J. 2000. Idiosyncratic risk: An empirical analysis, with implications for the risk of relative-value trading strategies. International Monetary Fund working paper. Washington, D.C.: International Monetary Fund.
Roll, R. 1988. R^2. *Journal of Finance* 43 (July): 541–66.
Ross, S. A. 1987. The interrelations of finance and economics: Theoretical perspectives. *American Economic Review* 77 (May): 29–34.
Scholes, M. S. 1972. The market for securities: Substitution versus price pressure and the effects of information on share price. *Journal of Business* 45 (April): 179–211.
Shleifer, A. 1986. Do demand curves for stocks slope down? *Journal of Finance* 41 (July): 579–90.
Shleifer, A. 2000. *Inefficient markets: An introduction to behavioral finance.* Oxford: Oxford University Press.
Shleifer, A., and Vishny, R. 1997. The limits of arbitrage. *Journal of Finance* 52 (March): 35–55.

Spiro, L. N. 1998. How Long-Term rocked stocks, too. *Business Week* (November 9).
Stoll, H. R. 2000. Friction. *Journal of Finance* 55 (August): 1479–1514.
The U.S. Offshore Funds Directory. 1999. New York: The U.S. Offshore Hedge Funds Directory, Inc.